10-019482

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2005, 09:35:34 ; Search time 39 Seconds
(without alignments)
61.677 Million cell updates/sec

Title: SEQ1
Perfect score: 105
Sequence: 1 axaeaaekaakyaaeaaekaakaxa 25

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	71	67.6	924	2	T06636	hypothetical prote
2	65	61.9	168	2	T34804	hypothetical prote
3	63	60.0	179	2	F97683	50S ribosomal prot
4	63	60.0	179	2	AF2908	50S ribosomal prot
5	63	60.0	421	2	JV0057	tolA protein - Esc
6	60	57.1	177	2	E87294	ATP synthase F0, B
7	58	55.2	354	1	GNVVSR	genome polyprotein
8	58	55.2	375	2	A71625	rifin PFB0035c - m
9	57	54.3	394	2	F90725	membrane spanning
10	57	54.3	394	2	G85576	membrane spanning
11	57	54.3	909	2	T06635	hypothetical prote
12	56	53.3	101	2	H59099	hypothetical prote
13	56	53.3	228	2	E87612	cytochrome c, memb
14	56	53.3	347	2	E83525	TolA protein PA097
15	56	53.3	356	2	A82152	tolA protein VC183
16	56	53.3	564	2	AH2328	ATP-binding protei
17	56	53.3	592	1	IKEBCA	colicin A - Citrob
18	55	52.4	97	2	S02376	antifreeze protein
19	55	52.4	110	2	T37490	ribosomal protein
20	55	52.4	555	2	S04909	embryonic protein
21	55	52.4	1110	2	I51116	NF-180 - sea lampr
22	55	52.4	1147	2	T35781	hypothetical prote
23	54.5	51.9	1203	2	C95229	DNA-directed RNA p
24	54.5	51.9	1216	2	G98093	DNA-directed RNA p
25	54	51.4	217	2	A26721	histone H1-gamma,
26	54	51.4	310	2	T34809	ribosomal protein
27	54	51.4	643	1	T07064	seed biotin-contai
28	54	51.4	1156	2	T34852	probable secreted
29	54	51.4	4687	1	A39638	plectin - rat
30	53.5	51.0	45	2	A05163	antifreeze protein
31	53.5	51.0	846	2	S52418	GTP-binding regula
32	53	50.5	40	1	FDFI8G	antifreeze protein
33	53	50.5	205	2	S19114	cgcr-1 protein - C
34	53	50.5	376	2	AG0592	tolA protein [impo
35	53	50.5	388	2	AC0138	TolA colicin impor
36	53	50.5	893	2	T38147	dolichyl-phosphate
37	52.5	50.0	1175	2	H83437	hypothetical prote
38	52	49.5	97	2	G60110	repetitive protein
39	52	49.5	147	2	D86389	hypothetical prote
40	52	49.5	192	2	T26386	hypothetical prote
41	52	49.5	210	2	A25550	histone H1 - sea u
42	52	49.5	248	1	HSUR1P	histone H1, gonada
43	52	49.5	288	2	S58219	ABA-inducible prot
44	52	49.5	294	2	S32234	transcription anti
45	52	49.5	294	2	S41061	probable transcrip

ALIGNMENTS

RESULT 1
T06636
hypothetical protein T20K18.130 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06636
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew
submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15790
A;Accession: T06636
A;Molecule type: DNA
A;Residues: 1-924 <BEV>
A;Cross-references: UNIPROT:Q9SU08; EMBL:AL049640; GSPDB:GN00062; ATSP:T20K18.130
A;Experimental source: cultivar Columbia; BAC clone T20K18
C;Genetics:
A;Gene: ATSP:T20K18.130
A;Map position: 4
A;Introns: 209/2; 699/3; 753/3; 785/2; 807/2; 853/3; 912/3

Query Match 67.6%; Score 71; DB 2; Length 924;
Best Local Similarity 68.0%; Pred. No. 0.71;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| | :|||| |||| ||| ||:|
Db        603 AAAGARDKAAKAAAEAREKAEKAAA 627
```

RESULT 2
T34804
hypothetical protein SC2E1.36 SC2E1.36 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T34804
R;Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1998
A;Reference number: Z21557
A;Accession: T34804
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-168 <MUR>
A;Cross-references: UNIPROT:O69907; EMBL:AL023797; PIDN:CAA19411.1; GSPDB:GN00070; SCOED
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC2E1.36

Query Match 61.9%; Score 65; DB 2; Length 168;
Best Local Similarity 62.5%; Pred. No. 0.81;
Matches 15; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:|:||||| : || ||:||| |:
```

```
Db        106 AEAKAAEKAERAAARAAKKAAPAP 129


RESULT 3
F97683
50S ribosomal protein L19 [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: F97683
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: F97683
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-179 <KUR>
A;Cross-references: UNIPROT:Q8UBZ5; GB:AE007869; PIDN:AAK88423.1; PID:g15157917; GSPDB:G
C;Genetics:
A;Gene: AGR_C_4900
A;Map position: circular chromosome
C;Superfamily: Escherichia coli ribosomal protein L19

  Query Match             60.0%;  Score 63;  DB 2;  Length 179;
  Best Local Similarity   69.2%;  Pred. No. 1.4;
  Matches   18;  Conservative    3;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAE--AAEKAAKAX 24
              |:| ||||||  |||  |||:||||:
Db        149 AQALAAEKAAAEAAEAKAAEEAAKAA 174


RESULT 4
AF2908
50S ribosomal protein L19 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AF2908
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AF2908
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-179 <KUR>
A;Cross-references: UNIPROT:Q8UBZ5; GB:AE008688; PIDN:AAL43684.1; PID:g17741210; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: rplS
A;Map position: circular chromosome
C;Superfamily: Escherichia coli ribosomal protein L19

  Query Match             60.0%;  Score 63;  DB 2;  Length 179;
  Best Local Similarity   69.2%;  Pred. No. 1.4;
  Matches   18;  Conservative    3;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAE--AAEKAAKAX 24
              |:| ||||||  |||  |||:||||:
Db        149 AQALAAEKAAAEAAEAKAAEEAAKAA 174


RESULT 5
JV0057
tolA protein - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: JV0057; B64810
R;Levengood, S.K.; Webster, R.E.
J. Bacteriol. 171, 6600-6609, 1989
A;Title: Nucleotide sequences of the tolA and tolB genes and localization of their produc
A;Reference number: JV0057; MUID:90078104; PMID:2687247
A;Accession: JV0057
A;Molecule type: DNA
A;Residues: 1-421 <LEV>
A;Cross-references: UNIPROT:P19934; GB:M28232; NID:g148018; PIDN:AAA24683.1; PID:g148019
A;Experimental source: strain JM105
A;Note: the authors translated the initiation codon GTG for residue 1 as Val
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: B64810
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-421 <BLAT>
A;Cross-references: GB:AE000177; GB:U00096; NID:g1786955; PIDN:AAC73833.1; PID:g1786960;
A;Experimental source: strain K-12, substrain MG1655
C;Comment: tolA and tolB proteins are necessary for colicins E2, E3, A, and K to reach th
C;Genetics:
A;Gene: tolA
A;Map position: 17 min
A;Start codon: GTG
C;Keywords: nucleotide binding; P-loop; transmembrane protein
F;14-34/Domain: transmembrane #status predicted <MSS>
F;78-301/Domain: helical #status predicted <HSR>
F;355-362/Region: nucleotide-binding motif A (P-loop)

  Query Match             60.0%;  Score 63;  DB 2;  Length 421;
  Best Local Similarity   60.0%;  Pred. No. 3;
  Matches   15;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|:|||:||| ||  |:| |:|:|
Db        151 ADAKAAEEAAKKAAADAKKKAEAEA 175


RESULT 6
E87294
ATP synthase F0, B' subunit [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: E87294
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.E
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolona
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: E87294
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-177 <STO>
A;Cross-references: UNIPROT:Q9AB65; GB:AE005673; NID:g13421521; PIDN:AAK22353.1; GSPDB:GN
C;Genetics:
A;Gene: CC0366

  Query Match             57.1%;  Score 60;  DB 2;  Length 177;
  Best Local Similarity   60.0%;  Pred. No. 3.2;
  Matches   15;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|||||: ||  |  ||| | |:|
Db        110 ASAEAAERQAKEEAVLAEKLAAAEA 134


RESULT 7
GNVVSR
genome polyprotein 1 - tomato ringspot virus (strain raspberry) (fragment)
C;Species: tomato ringspot virus
```

```
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C;Accession: A40787
R;Rott, M.E.; Tremaine, J.H.; Rochon, D.M.
Virology 185, 468-472, 1991
A;Title: Comparison of the 5' and 3' termini of tomato ringspot virus RNA1 and RNA2: evi
A;Reference number: A40787; MUID:92024112; PMID:1926788
A;Accession: A40787
A;Molecule type: genomic RNA
A;Residues: 1-354 <ROT>
A;Cross-references: UNIPROT:P29150; GB:M73822; NID:g335267; PIDN:AAA47941.1; PID:g555406
C;Genetics:
A;Map position: segment 1
C;Superfamily: tomato ringspot virus genome polyprotein
C;Keywords: glycoprotein; polyprotein
F;270/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             55.2%;  Score 58;  DB 1;  Length 354;
  Best Local Similarity   70.0%;  Pred. No. 9.7;
  Matches   14;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          6 AEKAAKYAAEAAEKAAKAXA 25
              | ||||||| || | | |:|
Db        180 ARKAAKYAAFAARKKAAAVA 199


RESULT 8
A71625
rifin PFB0035c - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
C;Accession: A71625
R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
.; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A;Reference number: A71600; MUID:99021743; PMID:9804551
A;Accession: A71625
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-375 <GAR>
A;Cross-references: UNIPROT:O96113; GB:AE001367; GB:AE001362; NID:g3845074; PIDN:AAC7179
A;Experimental source: clone 3D7
C;Genetics:
A;Gene: PFB0035c
C;Superfamily: Plasmodium falciparum rifin PFB1005w

  Query Match             55.2%;  Score 58;  DB 2;  Length 375;
  Best Local Similarity   56.5%;  Pred. No. 10;
  Matches   13;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              : | ||:||| ||:|||| |:
Db        292 IVEGAEQAAKAAAKAAEKGVTAA 314


RESULT 9
F90725
membrane spanning protein TolA [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F90725
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90725
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-394 <HAY>
A;Cross-references: UNIPROT:Q8X965; GB:BA000007; PIDN:BAB34197.1; PID:g13360233; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs0774

  Query Match             54.3%;  Score 57;  DB 2;  Length 394;
  Best Local Similarity   56.0%;  Pred. No. 14;
  Matches   14;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |: :|||:||| ||  |:| |:|:|
Db        151 ADDKAAEEAAKKAAADAKKKAEAEA 175


RESULT 10
G85576
membrane spanning protein TolA [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: G85576
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: G85576
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-394 <STO>
A;Cross-references: UNIPROT:Q8X965; GB:AE005174; NID:g12513672; PIDN:AAG55075.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: tolA

  Query Match             54.3%;  Score 57;  DB 2;  Length 394;
  Best Local Similarity   56.0%;  Pred. No. 14;
  Matches   14;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |: :|||:||| ||  |:| |:|:|
Db        151 ADDKAAEEAAKKAAADAKKKAEAEA 175


RESULT 11
T06635
hypothetical protein T20K18.120 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06635
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew
submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15790
A;Accession: T06635
A;Molecule type: DNA
A;Residues: 1-909 <BEV>
A;Cross-references: UNIPROT:Q9SU09; EMBL:AL049640; GSPDB:GN00062; ATSP:T20K18.120
A;Experimental source: cultivar Columbia; BAC clone T20K18
C;Genetics:
A;Gene: ATSP:T20K18.120
A;Map position: 4
A;Introns: 205/2; 686/3; 740/3; 772/2; 808/3; 838/3; 897/3

  Query Match             54.3%;  Score 57;  DB 2;  Length 909;
  Best Local Similarity   66.7%;  Pred. No. 28;
  Matches   14;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAA 21
              |:||| |:||  | | |||||
Db        593 AHAEARERAAAGAREKAEKAA 613


RESULT 12
H59099
hypothetical protein pXO1-72 - Bacillus anthracis virulence plasmid pXO1
C;Species: Bacillus anthracis
C;Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 09-Jul-2004
```

C;Accession: H59099
R;Okinaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler
J. Bacteriol. 181, 6509-6515, 1999
A;Title: Sequence and organization of pXO1, the large Bacillus anthracis plasmid harbori
A;Reference number: A59091; MUID:99445483; PMID:10515943
A;Accession: H59099
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-101 <OKI>
A;Cross-references: UNIPROT:Q9X342; GB:AF065404; NID:g4894216; PIDN:AAD32376.1; PID:g489
A;Experimental source: strain Sterne
A;Note: similar to hypothetical, locus Clo tetP Clostridium perfringens (L20800)
C;Genetics:
A;Gene: pXO1-72
A;Genome: plasmid
C;Superfamily: Bacillus anthracis virulence plasmid pXO1 hypothetical protein pXO1-72

  Query Match 53.3%; Score 56; DB 2; Length 101;
  Best Local Similarity 63.6%; Pred. No. 5.6;
  Matches 14; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AXAEAAEKAAKYAAEAAEKAAK 22
|: :||||||| ||| | | |
Db 44 AEEKAAEKAAKEAAEEARIATK 65

RESULT 13
E87612
cytochrome c, membrane-bound [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 15-Mar-2004
C;Accession: E87612
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: E87612
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-228 <STO>
A;Cross-references: GB:AE005673; NID:g13424561; PIDN:AAK24897.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC2935
C;Superfamily: membrane-bound c-type cytochrome; cytochrome c homology
C;Keywords: chromoprotein; heme; iron; metalloprotein
F;81,84/Binding site: heme (Cys) (covalent) #status predicted
F;85/Binding site: heme iron (His) (axial ligand) #status predicted
F;150/Binding site: heme iron (Met) (axial ligand) #status predicted

  Query Match 53.3%; Score 56; DB 2; Length 228;
  Best Local Similarity 60.0%; Pred. No. 11;
  Matches 15; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
|:|| | ||: || ||| || |:|
Db 187 APAEGAAPAAEGAAPAAEGAAPAAA 211

RESULT 14
E83525
TolA protein PA0971 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: E83525
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83525
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-347 <STO>
A;Cross-references: UNIPROT:P50600; GB:AE004530; GB:AE004091; NID:g9946865; PIDN:AAG0436
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: tolA; PA0971

  Query Match 53.3%; Score 56; DB 2; Length 347;
  Best Local Similarity 56.0%; Pred. No. 16;
  Matches 14; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
|: :||| | | ||| |:| | |:|
Db 171 AKKKAAEDAKKKAAEEAKKKAAAEA 195

RESULT 15
A82152
tolA protein VC1837 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: A82152
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: A82152
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-356 <HEI>
A;Cross-references: UNIPROT:Q9KR10; GB:AE004259; GB:AE003852; NID:g9656353; PIDN:AAF9498
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC1837
A;Map position: 1

  Query Match 53.3%; Score 56; DB 2; Length 356;
  Best Local Similarity 54.5%; Pred. No. 17;
  Matches 18; Conservative 4; Mismatches 3; Indels 8; Gaps 2;

Qy 1 AXAE----AAEKAAKYAAEAA----EKAAKAXA 25
|:|| | ||||| ||: | |:||||:|
Db 199 AKAEQERLAKEKAAKEAADKAKKEKERAAKAEA 231

RESULT 16
AH2328
ATP-binding protein of ABC transporter all4183 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AH2328
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anab
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2328
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-564 <KUR>
A;Cross-references: UNIPROT:Q8YPL1; GB:BA000019; PIDN:BAB75882.1; PID:g17133318; GSPDB:GN
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all4183
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

  Query Match 53.3%; Score 56; DB 2; Length 564;
  Best Local Similarity 61.9%; Pred. No. 25;
  Matches 13; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```
Qy          2 XAEAAEKAAKYAAEAAEKAAK 22
              :| ||||||| ||::|: :||
Db        542 EAIAAEKAAKKAAKSAKSSAK 562


RESULT 17
IKEBCA
colicin A - Citrobacter freundii (strain CA31) plasmid ColA
C;Species: Citrobacter freundii
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C;Accession: I40784; A03504; I40777
R;Morlon, J.; Chartier, M.; Bidaud, M.; Lazdunski, C.
Mol. Gen. Genet. 211, 231-243, 1988
A;Title: The complete nucleotide sequence of the colicinogenic plasmid ColA. High extent
A;Reference number: I40778; MUID:88174422; PMID:2832701
A;Accession: I40784
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-592 <RES>
A;Cross-references: UNIPROT:P04480; GB:M37402; NID:g144661; PIDN:AAA72879.1; PID:g144667
A;Experimental source: plasmid ColA
R;Morlon, J.; Lloubes, R.; Varenne, S.; Chartier, M.; Lazdunski, C.
J. Mol. Biol. 170, 271-285, 1983
A;Title: Complete nucleotide sequence of the structural gene for colicin A, a gene trans
A;Reference number: A03504; MUID:84036205; PMID:6313941
A;Accession: A03504
A;Molecule type: DNA
A;Residues: 1-592 <MOR>
A;Cross-references: GB:X01008; GB:X00034; NID:g40459; PIDN:CAA25503.1; PID:g40460
R;Morlon, J.; Lloubes, R.; Chartier, M.; Bonicel, J.; Lazdunski, C.
EMBO J. 2, 787-789, 1983
A;Title: Nucleotide sequence of promoter, operator and amino-terminal region of caa, the
A;Reference number: I40777; MUID:84057757; PMID:6641715
A;Accession: I40777
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-53,'X',55-70 <RE2>
A;Cross-references: GB:M26369; NID:g144659; PIDN:AAA98057.1; PID:g144660
A;Experimental source: plasmid ColA
C;Comment: This protein acts to depolarize the bacterial inner membrane, most likely by
C;Genetics:
A;Gene: caa
A;Genome: plasmid
C;Superfamily: colicin IB
C;Keywords: antibiotic; bacteriocin; toxin; transmembrane protein

  Query Match             53.3%;  Score 56;  DB 1;  Length 592;
  Best Local Similarity   54.2%;  Pred. No. 26;
  Matches   13;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:|||||| :  || ||:  :|:
Db        364 AAAEAAEKARQRQAEEAERQRQAM 387


RESULT 18
S02376
antifreeze protein precursor - yellowtail flounder
C;Species: Limanda ferruginea (yellowtail flounder)
C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C;Accession: S02376
R;Scott, G.K.; Davies, P.L.; Shears, M.A.; Fletcher, G.L.
Eur. J. Biochem. 168, 629-633, 1987
A;Title: Structural variations in the alanine-rich antifreeze proteins of the Pleuronect
A;Reference number: S02376; MUID:88029483; PMID:3665937
A;Accession: S02376
A;Molecule type: mRNA
A;Residues: 1-97 <SCO>
A;Cross-references: UNIPROT:P09031; EMBL:X06356; NID:g64041; PIDN:CAA29655.1; PID:g64042
A;Note: part of this sequence, including the amino end of the mature protein, was confir
C;Superfamily: antifreeze protein
C;Keywords: antifreeze
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-48/Domain: propeptide #status predicted <PRO>
F;49-96/Product: antifreeze protein #status predicted <MAT>

  Query Match             52.4%;  Score 55;  DB 2;  Length 97;
  Best Local Similarity   56.0%;  Pred. No. 7.1;
  Matches   14;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|  |  ||| ||: |  ||||:|
Db         57 AAAATAAAAAKAAADTAAAAAKAAA 81


RESULT 19
T37490
ribosomal protein rpa6 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T37490
R;Bothe, G.; Pohl, T.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, November 1999
A;Reference number: Z21718
A;Accession: T37490
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-110 <BOT>
A;Cross-references: UNIPROT:O14317; EMBL:AL132769; PIDN:CAB59884.1; GSPDB:GN00066; SPDB:
A;Experimental source: strain 972h-; cosmid c1071
C;Genetics:
A;Gene: SPDB:SPAC1071.08
A;Map position: 1
A;Introns: 103/3
C;Superfamily: rat acidic ribosomal protein P1

  Query Match             52.4%;  Score 55;  DB 2;  Length 110;
  Best Local Similarity   60.0%;  Pred. No. 7.9;
  Matches   15;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||  ||  || |||:||| :|
Db         70 AAAAAAPAAAGGAAPAAEEAAKEEA 94


RESULT 20
S04909
embryonic protein DC8 (clone 8/10) - carrot
C;Species: Daucus carota (carrot)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: S04909
R;Franz, G.; Hatzopoulos, P.; Jones, T.J.; Krauss, M.; Sung, Z.R.
Mol. Gen. Genet. 218, 143-151, 1989
A;Title: Molecular and genetic analysis of an embryonic gene, DC 8, from Daucus carota L
A;Reference number: S04909; MUID:89384429; PMID:2571069
A;Accession: S04909
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-555 <FRA>
A;Cross-references: UNIPROT:P20075; GB:X16131; NID:g18333; PIDN:CAA34258.2; PID:g4902464
C;Superfamily: pea seed biotin-containing protein

  Query Match             52.4%;  Score 55;  DB 2;  Length 555;
  Best Local Similarity   52.2%;  Pred. No. 32;
  Matches   12;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :||| ||  :|   ||:|||:|:
Db        195 AAEAKEKTGEYKDYAAQKAAEAK 217


RESULT 21
I51116
NF-180 - sea lamprey
C;Species: Petromyzon marinus (sea lamprey)
```

```
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: I51116
R;Jacobs, A.J.; Kamholz, J.; Selzer, M.E.
Brain Res. Mol. Brain Res. 29, 43-52, 1995
A;Title: The single lamprey neurofilament subunit (NF-180) lacks multiphosphorylation re
A;Reference number: I51116; MUID:95287814; PMID:7770000
A;Accession: I51116
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1110 <JAC>
A;Cross-references: UNIPROT:Q91255; EMBL:U19361; NID:g632548; PIDN:AAA80106.1; PID:g6325
C;Superfamily: neurofilament triplet H protein

  Query Match             52.4%;  Score 55;  DB 2;  Length 1110;
  Best Local Similarity   65.0%;  Pred. No. 57;
  Matches   13;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKA 20
              |:||| |:|||   ||||:|
Db        725 AEAEAEEEAAKSEEEAAEEA 744


RESULT 22
T35781
hypothetical protein SC8A6.14c SC8A6.14c - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T35781
R;Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1998
A;Reference number: Z21570
A;Accession: T35781
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1147 <SEE>
A;Cross-references: UNIPROT:O87848; EMBL:AL031013; PIDN:CAA19786.1; GSPDB:GN00070; SCOED
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC8A6.14c

  Query Match             52.4%;  Score 55;  DB 2;  Length 1147;
  Best Local Similarity   50.0%;  Pred. No. 59;
  Matches   12;  Conservative    7;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:|:|||:  : :| || |||:|:
Db        264 AEAQAAEQDVQRSARAANKAARAA 287


RESULT 23
C95229
DNA-directed RNA polymerase, beta chain [imported] - Streptococcus pneumoniae (strain TI
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: C95229
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: C95229
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1203 <KUR>
A;Cross-references: UNIPROT:Q97NQ7; GB:AE005672; PIDN:AAK76028.1; PID:g14973467; GSPDB:G
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP1961
C;Superfamily: DNA-directed RNA polymerase beta chain

  Query Match             51.9%;  Score 54.5;  DB 2;  Length 1203;
  Best Local Similarity   53.8%;  Pred. No. 70;
  Matches   14;  Conservative    6;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 AXAEAAEKA-AKYAAEAAEKAAKAXA 25
              |: :||::| | : || |||| ||:|
Db       1170 AREKAAQEAKAAFEAEEAEKATKAEA 1195


RESULT 24
G98093
DNA-directed RNA polymerase (EC 2.7.7.6) [imported] - Streptococcus pneumoniae (strain R6
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: G98093
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczak, L.; Burgett, S.; DeHoff, B.S.; Es
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Mc
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: G98093
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1216 <KUR>
A;Cross-references: UNIPROT:Q8DNF0; GB:AE007317; PIDN:AAL00580.1; PID:g15459460; GSPDB:GN
C;Genetics:
A;Gene: rpoB
C;Superfamily: DNA-directed RNA polymerase beta chain
C;Keywords: nucleotidyltransferase

  Query Match             51.9%;  Score 54.5;  DB 2;  Length 1216;
  Best Local Similarity   53.8%;  Pred. No. 70;
  Matches   14;  Conservative    6;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 AXAEAAEKA-AKYAAEAAEKAAKAXA 25
              |: :||::| | : || |||| ||:|
Db       1183 AREKAAQEAKAAFEAEEAEKATKAEA 1208


RESULT 25
A26721
histone H1-gamma, embryonic - sea urchin (Strongylocentrotus purpuratus)
C;Species: Strongylocentrotus purpuratus (purple urchin)
C;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 09-Jul-2004
C;Accession: A26721
R;Knowles, J.A.; Lai, Z.C.; Childs, G.J.
Mol. Cell. Biol. 7, 478-485, 1987
A;Title: Isolation, characterization, and expression of the gene encoding the late histor
A;Reference number: A26721; MUID:87172742; PMID:3031476
A;Accession: A26721
A;Molecule type: DNA
A;Residues: 1-217 <KNO>
A;Cross-references: UNIPROT:P07796; GB:M16033; NID:g161517; PIDN:AAA30059.1; PID:g161518
C;Superfamily: histone H1
C;Keywords: chromosomal protein; DNA binding; nucleosome; nucleus
F;2-217/Product: histone H1-gamma, embryonic #status predicted <MAT>

  Query Match             51.4%;  Score 54;  DB 2;  Length 217;
  Best Local Similarity   54.2%;  Pred. No. 18;
  Matches   13;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:|: |:|||| | :||: | ||:
Db        189 AAAKPAKKAAKPAKKAAKPAKKAA 212


RESULT 26
T34809
ribosomal protein S2 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T34809
```

```
R;Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1998
A;Reference number: Z21557
A;Accession: T34809
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-310 <MUR>
A;Cross-references: UNIPROT:O31212; EMBL:AL023797; PIDN:CAA19416.1; GSPDB:GN00070; SCOED
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: rpsB; SCOEDB:SC2E1.41
C;Superfamily: Escherichia coli ribosomal protein S2

  Query Match             51.4%;  Score 54;  DB 2;  Length 310;
  Best Local Similarity   58.3%;  Pred. No. 25;
  Matches   14;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :|||||  |  || | | ||:|:|
Db        281 AAEAAEAPAAEAAPAEEPAAEAPA 304


RESULT 27
T07064
seed biotin-containing protein LEA [validated] - soybean
C;Species: Glycine max (soybean)
C;Date: 01-Sep-2000 #sequence_revision 01-Sep-2000 #text_change 09-Jul-2004
C;Accession: T07064
R;Hsing, Y.C.; Tsou, C.H.; Hsu, T.F.; Chen, Z.Y.; Hsieh, K.L.; Hsieh, J.S.; Chow, T.Y.
Plant Mol. Biol. 38, 481-490, 1998
A;Title: Tissue- and stage-specific expression of a soybean (Glycine max L.) seed-matura
A;Reference number: Z15895; MUID:98418627; PMID:9747855
A;Accession: T07064
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-643 <HSI>
A;Cross-references: UNIPROT:Q39846; EMBL:U59626; NID:g1389896; PIDN:AAC61783.1; PID:g138
A;Experimental source: strain Shi-Shi; cotyledon
C;Superfamily: pea seed biotin-containing protein
C;Keywords: biotin binding; seed
F;125/Binding site: biotin (Lys) (covalent) #status predicted

  Query Match             51.4%;  Score 54;  DB 1;  Length 643;
  Best Local Similarity   52.2%;  Pred. No. 47;
  Matches   12;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|  ||||  |  :|||||  |:
Db        302 TAPVAEKAKDYTLQAAEKAKSAG 324


RESULT 28
T34852
probable secreted protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T34852
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1999
A;Reference number: Z21559
A;Accession: T34852
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1156 <OLI>
A;Cross-references: UNIPROT:Q9Z5A4; EMBL:AL035478; PIDN:CAB36606.1; GSPDB:GN00070; SCOED
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC2G5.19

  Query Match             51.4%;  Score 54;  DB 2;  Length 1156;
  Best Local Similarity   58.3%;  Pred. No. 77;
  Matches   14;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              ::| || |||  ||  ||||| |:
Db        473 SAAAAARKAADSAAAHAEKAADAA 496


RESULT 29
A39638
plectin - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A39638; S21876
R;Wiche, G.; Becker, B.; Luber, K.; Weitzer, G.; Castanon, M.J.; Hauptmann, R.; Stratowa
J. Cell Biol. 114, 83-99, 1991
A;Title: Cloning and sequencing of rat plectin indicates a 466-kD polypeptide chain with
A;Reference number: A39638; MUID:91268156; PMID:2050743
A;Accession: A39638
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-4687 <WIC>
A;Cross-references: UNIPROT:P30427; EMBL:X59601; NID:g1292885; PIDN:CAA42169.1; PID:g156
C;Superfamily: plectin; alpha-actinin actin-binding domain homology; ribosomal protein S
C;Keywords: cytoskeleton; transmembrane protein
F;6-103/Domain: ribosomal protein S10 homology <RS10>
F;184-399/Domain: alpha-actinin actin-binding domain homology <ACT>

  Query Match             51.4%;  Score 54;  DB 1;  Length 4687;
  Best Local Similarity   48.0%;  Pred. No. 2.6e+02;
  Matches   12;  Conservative    7;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              ::|||| :||: | || |:| : :|
Db       2221 SEAEAARRAAEEAEEAREQAEREAA 2245


RESULT 30
A05163
antifreeze protein SS-8 - shorthorn sculpin
C;Species: Myoxocephalus scorpius (shorthorn sculpin, daddy sculpin)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 09-Jul-2004
C;Accession: A05163
R;Hew, C.L.; Joshi, S.; Wang, N.C.; Kao, M.H.; Ananthanarayanan, V.S.
Eur. J. Biochem. 151, 167-172, 1985
A;Title: Structures of shorthorn sculpin antifreeze polypeptides.
A;Reference number: A91150; MUID:85285003; PMID:4029130
A;Accession: A05163
A;Molecule type: protein
A;Residues: 1-45 <HEW>
A;Cross-references: UNIPROT:P04368
C;Superfamily: antifreeze protein
C;Keywords: antifreeze; blocked amino end; plasma
F;9-45/Region: alanine-rich
F;1/Modified site: blocked amino end (Met) #status experimental

  Query Match             51.0%;  Score 53.5;  DB 2;  Length 45;
  Best Local Similarity   64.0%;  Pred. No. 5.5;
  Matches   16;  Conservative    3;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 XAEAAEKAAKYAAEAAEK-AAKAXA 25
              :| ||  ||| ||:|| | ||||:|
Db         13 LAAAAALAAKTAADAAAKAAAKAAA 37


RESULT 31
S52418
GTP-binding regulatory protein Gs alpha-XL chain - rat
N;Alternate names: G protein XL-alpha-s
C;Species: Rattus norvegicus (Norway rat)
C;Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 02-Feb-2001
C;Accession: S52418
R;Kehlenbach, R.H.; Matthey, J.; Huttner, W.B.
Nature 372, 804-809, 1994
A;Title: XL-alpha-s is a new type of G protein.
A;Reference number: S52418; MUID:95089824; PMID:7997272
```

```
A;Accession: S52418
A;Molecule type: mRNA
A;Residues: 1-846 <KEH>
A;Cross-references: EMBL:X84047; NID:g642267; PIDN:CAA58866.1; PID:g642268
R;Kehlenbach, R.H.; Matthey, J.; Huttner, W.B.
Nature 375, 253, 1995
A;Title: Correction: XLalphas is a new type of G protein.
A;Reference number: S58911
A;Contents: annotation; assignment of start_codon
A;Note: experimental data from this paper suggest that the translation is initiated at p
C;Keywords: GTP binding; nucleotide binding; P-loop; signal transduction
F;132-846/Product: GTP-binding regulatory protein Gs alpha-XL chain #status experimental
F;499-506/Region: nucleotide-binding motif A (P-loop)
F;744-747/Region: GTP-binding NKXD motif

  Query Match             51.0%;  Score 53.5;  DB 2;  Length 846;
  Best Local Similarity   61.5%;  Pred. No. 67;
  Matches   16;  Conservative    4;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAE-AAEKAAKAXA 25
              |:| ||| ||: ||| ||| ||: :|
Db        170 AAAAAAEPAAEPAAEPAAEPAAEPAA 195


RESULT 32
FDFI8G
antifreeze protein GS-8 - grubby sculpin
C;Species: Myoxocephalus aenaeus (grubby sculpin)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C;Accession: S07046
R;Chakrabartty, A.; Hew, C.L.; Shears, M.; Fletcher, G.
Can. J. Zool. 66, 403-408, 1988
A;Title: Primary structures of the alanine-rich antifreeze polypeptides from grubby scul
A;Reference number: S06417
A;Accession: S07046
A;Molecule type: protein
A;Residues: 1-40 <CHA>
A;Cross-references: UNIPROT:P20617
C;Superfamily: antifreeze protein
C;Keywords: antifreeze; blocked amino end
F;1/Modified site: blocked amino end (Met) #status experimental

  Query Match             50.5%;  Score 53;  DB 1;  Length 40;
  Best Local Similarity   58.3%;  Pred. No. 5.6;
  Matches   14;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||  ||| ||:|| |||  :|
Db         14 AAAAAALAAKTAADAAAKAAAIAA 37


RESULT 33
S19114
cgcr-1 protein - Chlamydomonas reinhardtii (fragment)
C;Species: Chlamydomonas reinhardtii
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S19114
R;Wakarchuk, W.W.; Mueller, F.W.; Beck, C.F.
Plant Mol. Biol. 18, 143-146, 1992
A;Title: Two GC-rich DNA elements of Chlamydomonas reinhardtii with complex arrangements
A;Reference number: S19113; MUID:92119224; PMID:1731966
A;Accession: S19114
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-205 <WAK>
A;Cross-references: UNIPROT:Q39597; EMBL:X17207
C;Superfamily: phage lambda hypothetical protein 401

  Query Match             50.5%;  Score 53;  DB 2;  Length 205;
  Best Local Similarity   52.0%;  Pred. No. 23;
  Matches   13;  Conservative    6;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|:|| :||   | ||: ||:|:|
Db        124 AAAQAAAQAAAERAAAAQAAAQAAA 148


RESULT 34
AG0592
tolA protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain CT18)
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AG0592
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serova
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AG0592
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-376 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05209.1; PID:g16501979; GSPDB:GN00176
C;Genetics:
A;Gene: STY0793

  Query Match             50.5%;  Score 53;  DB 2;  Length 376;
  Best Local Similarity   52.0%;  Pred. No. 38;
  Matches   13;  Conservative    6;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:||||: ||:   :|  :||||:|
Db        177 AEAEAAKAAAEAKKKAEAEAAKAAA 201


RESULT 35
AC0138
TolA colicin import membrane protein [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC0138
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.,
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; P
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, P
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AC0138
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-388 <KUR>
A;Cross-references: UNIPROT:Q8ZGZ2; GB:AL590842; PIDN:CAC89966.1; PID:g15979190; GSPDB:GN
C;Genetics:
A;Gene: tolA

  Query Match             50.5%;  Score 53;  DB 2;  Length 388;
  Best Local Similarity   59.1%;  Pred. No. 39;
  Matches   13;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 EAAEKAAKYAAEAAEKAAKAXA 25
              :| | | | ||:|||| | |:|
Db        212 KAVEVAEKAAADAAEKKAAADA 233


RESULT 36
T38147
dolichyl-phosphate-mannose-protein mannosyl transferase - fission yeast (Schizosaccharomy
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T38147
R;Pearson, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z21774
```

```
A;Accession: T38147
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-893 <PEA>
A;Cross-references: UNIPROT:O13898; EMBL:Z99295; PIDN:CAB16577.1; GSPDB:GN00066; SPDB:SP
A;Experimental source: strain 972h-; cosmid c22A12
C;Genetics:
A;Gene: SPDB:SPAC22A12.07c
A;Map position: 1
C;Superfamily: dolichyl-phosphate-mannose-protein mannosyltransferase

  Query Match             50.5%;  Score 53;  DB 2;  Length 893;
  Best Local Similarity   48.0%;  Pred. No. 80;
  Matches   12;  Conservative    6;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |: || | | | |:||||:::  :|
Db        786 AEQEAREAAEKAASEAAERSSSEAA 810


RESULT 37
H83437
hypothetical protein PA1669 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: H83437
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathc
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83437
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1175 <STO>
A;Cross-references: UNIPROT:Q9I356; GB:AE004594; GB:AE004091; NID:g9947630; PIDN:AAG050S
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA1669

  Query Match             50.0%;  Score 52.5;  DB 2;  Length 1175;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+02;
  Matches   16;  Conservative    4;  Mismatches    5;  Indels    7;  Gaps    1;

Qy          1 AXAEAAEKAA-------KYAAEAAEKAAKAXA 25
              |:|||::|||       | || || ||  |:|
Db        802 ALAEASDKAAEKGGKLGKAAAAAAGKARDALA 833


RESULT 38
G60110
repetitive protein antigen 69/70 - Trypanosoma cruzi (fragment)
C;Species: Trypanosoma cruzi
C;Date: 10-Nov-1992 #sequence_revision 10-Nov-1992 #text_change 09-Jul-2004
C;Accession: G60110
R;Hoft, D.F.; Kim, K.S.; Otsu, K.; Moser, D.R.; Yost, W.J.; Blumin, J.H.; Donelson, J.E.
Infect. Immun. 57, 1959-1967, 1989
A;Title: Trypanosoma cruzi expresses diverse repetitive protein antigens.
A;Reference number: A60110; MUID:89277508; PMID:2659529
A;Accession: G60110
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-97 <HOF>
A;Cross-references: UNIPROT:Q7M3W1
C;Superfamily: varicella-zoster virus gene 22 protein
C;Keywords: tandem repeat
F;1-85/Region: 7-residue repeats

  Query Match             49.5%;  Score 52;  DB 2;  Length 97;
  Best Local Similarity   56.0%;  Pred. No. 16;
  Matches   14;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|:||   || || |  ||| |:|
Db         51 APAKAAAAPAKTAAAAPAKAAAAPA 75


RESULT 39
D86389
hypothetical protein F28B23.4 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: D86389
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D86389
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-147 <STO>
A;Cross-references: UNIPROT:Q9C674; GB:AE005172; NID:g11079518; PIDN:AAG29229.1; GSPDB:G
C;Genetics:
A;Map position: 1

  Query Match             49.5%;  Score 52;  DB 2;  Length 147;
  Best Local Similarity   52.6%;  Pred. No. 22;
  Matches   10;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKA 20
              :|| |:|: ::||||::||
Db         61 VAETAKKSKEFAAEASKKA 79


RESULT 40
T26386
hypothetical protein Y105C5B.j - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T26386
R;McMurray, A.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z20208
A;Accession: T26386
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-192 <WIL>
A;Cross-references: UNIPROT:Q9NAM6; EMBL:AL110479; NID:e1542153; PIDN:CAB54358.1; CESP:Y
A;Experimental source: clone Y105C5B
C;Genetics:
A;Gene: CESP:Y105C5B.j
C;Superfamily: human S-phase kinase-associated protein 1A

  Query Match             49.5%;  Score 52;  DB 2;  Length 192;
  Best Local Similarity   70.6%;  Pred. No. 28;
  Matches   12;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 EAAEKAAKYAAEAAEKA 20
              ||||:|||  ||| :||
Db        163 EAAERAAKEKAEAEKKA 179


RESULT 41
A25550
histone H1 - sea urchin (Lytechinus pictus)
C;Species: Lytechinus pictus (painted urchin)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C;Accession: A25550
```

```
R;Knowles, J.A.; Childs, G.J.
Nucleic Acids Res. 14, 8121-8133, 1986
A;Title: Comparison of the late H1 histone genes of the sea urchins Lytechinus pictus ar
A;Reference number: A25550; MUID:87040778; PMID:3022245
A;Accession: A25550
A;Molecule type: DNA
A;Residues: 1-210 <KNO>
A;Cross-references: UNIPROT:P06144; GB:X04488; NID:g9616; PIDN:CAA28177.1; PID:g9617
C;Superfamily: histone H1
C;Keywords: chromosomal protein; DNA binding; nucleosome; nucleus

  Query Match             49.5%;  Score 52;  DB 2;  Length 210;
  Best Local Similarity   52.0%;  Pred. No. 30;
  Matches   13;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:  | : ||| ||: |:|||| :|
Db        179 AKKAAKKPAAKKAAKPAKKAAKKPA 203


RESULT 42
HSUR1P
histone H1, gonadal - sea urchin (Parechinus angulosus)
C;Species: Parechinus angulosus (angulate urchin)
C;Date: 31-Mar-1980 #sequence_revision 31-Mar-1980 #text_change 09-Jul-2004
C;Accession: A91090; A91091; A02586
R;Strickland, W.N.; Strickland, M.; de Groot, P.C.; von Holt, C.; Wittmann-Liebold, B.
Eur. J. Biochem. 104, 559-566, 1980
A;Title: The primary structure of histone H1 from sperm of the sea urchin Parechinus ang
rminal cyanogen bromide peptides.
A;Reference number: A91090; MUID:80156831; PMID:6767609
A;Contents: sequence of residues 1-84
A;Accession: A91090
A;Molecule type: protein
A;Residues: 1-248 <STR>
A;Cross-references: UNIPROT:P02256
R;Strickland, W.N.; Strickland, M.; Brandt, W.F.; von Holt, C.; Lehmann, A.; Wittmann-Li
Eur. J. Biochem. 104, 567-578, 1980
A;Title: The primary structure of histone H1 from sperm of the sea urchin Parechinus ang
A;Reference number: A91091; MUID:80156832; PMID:7363905
A;Accession: A91091
A;Molecule type: protein
A;Residues: 80-248 <ST2>
A;Note: 144-Arg was also found
C;Superfamily: histone H1
C;Keywords: DNA binding; nucleosome; sperm

  Query Match             49.5%;  Score 52;  DB 1;  Length 248;
  Best Local Similarity   54.2%;  Pred. No. 35;
  Matches   13;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :  || || |  | ||:|| ||:|
Db        124 KTSAAAKAKKAKAAAAKKARKAKA 147


RESULT 43
S58219
ABA-inducible protein, landform-specific - Riccia fluitans
C;Species: Riccia fluitans
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S58219
R;Hellwege, E.M.; Dietz, K.J.; Hartung, W.
submitted to the EMBL Data Library, July 1995
A;Description: Abscisic acid causes changes in gene expression involved in the induction
A;Reference number: S58219
A;Accession: S58219
A;Molecule type: mRNA
A;Residues: 1-288 <HEL>
A;Cross-references: UNIPROT:Q41154; EMBL:X89041; NID:g929818; PIDN:CAA61439.1; PID:g9298

  Query Match             49.5%;  Score 52;  DB 2;  Length 288;
  Best Local Similarity   45.8%;  Pred. No. 40;
  Matches   11;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :|| ||:|  | ||  |:|  |::
Db         75 GAEKAEQAKNYGAEKTEQAKSAAS 98


RESULT 44
S32234
transcription antitermination factor nusG - Streptomyces griseus
C;Species: Streptomyces griseus
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 09-Jul-2004
C;Accession: S32234
R;Kuester, K.; Kuberski, S.; Piepersberg, W.; Distler, J.
submitted to the EMBL Data Library, March 1993
A;Description: Cloning and nucleotide sequence analysis of the nusG-rplK-rplA-rplJ-rplL g
A;Reference number: S32234
A;Accession: S32234
A;Molecule type: DNA
A;Residues: 1-294 <KUE>
A;Cross-references: UNIPROT:P36260; EMBL:X72787; NID:g575399; PIDN:CAA51296.1; PID:g60358
C;Genetics:
A;Gene: nusG
A;Start codon: GTG
C;Superfamily: transcription antitermination factor nusG
C;Keywords: transcription antitermination; transcription factor

  Query Match             49.5%;  Score 52;  DB 2;  Length 294;
  Best Local Similarity   61.9%;  Pred. No. 40;
  Matches   13;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 EAAEKAAKYAAEAAEKAAKAX 24
              || |||| |||| | | |:
Db        212 EAEEKAAREAAEAEGKPAPAR 232


RESULT 45
S41061
probable transcription antitermination factor nusG - Streptomyces griseus (strain IFO133
C;Species: Streptomyces griseus
A;Variety: strain IFO13350
C;Date: 19-Mar-1997 #sequence_revision 12-Dec-1997 #text_change 09-Jul-2004
C;Accession: S41061
R;Miyake, K.; Onaka, H.; Horinouchi, S.; Beppu, T.
Biochim. Biophys. Acta 1217, 97-100, 1994
A;Title: Organization and nucleotide sequence of the secE-nusG region of Streptomyces gri
A;Reference number: S41059; MUID:94114580; PMID:8286423
A;Accession: S41061
A;Molecule type: DNA
A;Residues: 1-294 <MIY>
A;Cross-references: UNIPROT:P36260; EMBL:D17464; NID:g436786; PIDN:BAA04281.1; PID:g4838
A;Experimental source: strain IFO13350
C;Genetics:
A;Gene: nusG
A;Start codon: GTG
C;Function:
A;Description: may be involved in antibiotics production
C;Superfamily: transcription antitermination factor nusG

  Query Match             49.5%;  Score 52;  DB 2;  Length 294;
  Best Local Similarity   61.9%;  Pred. No. 40;
  Matches   13;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 EAAEKAAKYAAEAAEKAAKAX 24
              || |||| |||| | | |:
Db        212 EAEEKAAREAAEAEGKPAPAR 232


Search completed: July 11, 2005, 09:47:10
Job time : 42 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2005, 09:24:04 ; Search time 167 Seconds
(without alignments)
76.659 Million cell updates/sec

Title: SEQ1
Perfect score: 105
Sequence: 1 axaeaaekaakyaaeaaekaakaxa 25

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	71	67.6	485	2	Q8RXD0	Q8rxd0 arabidopsis
2	71	67.6	924	2	Q9SU08	Q9su08 arabidopsis
3	67	63.8	278	2	Q7Q0M9	Q7q0m9 anopheles g
4	65	61.9	168	2	O69907	streptomyce
5	63	60.0	179	1	RL19_AGRT5	Q8ubz5 agrobacteri
6	63	60.0	413	2	Q83SA1	Q83sa1 shigella fl
7	63	60.0	421	1	TOLA_ECOLI	P19934 escherichia
8	63	60.0	421	2	Q8FJT1	Q8fjt1 escherichia
9	61	58.1	441	2	Q6N8X8	Q6n8x8 rhodopseudo
10	61	58.1	593	2	Q8ZNE5	Q8zne5 salmonella
11	60.5	57.6	711	2	Q7Z3F5	Q7z3f5 homo sapien
12	60.5	57.6	730	1	ELS_HUMAN	P15502 homo sapien
13	60.5	57.6	757	2	Q14234	Q14234 homo sapien
14	60.5	57.6	757	2	Q75MU5	Q75mu5 homo sapien
15	60	57.1	105	2	Q6N503	Q6n503 rhodopseudo
16	60	57.1	177	2	Q9AB65	Q9ab65 caulobacter
17	60	57.1	371	2	Q6N4V4	Q6n4v4 rhodopseudo
18	60	57.1	572	2	Q6NH65	Q6nh65 corynebacte
19	60	57.1	738	2	Q6UBQ3	Q6ubq3 chlamydomon
20	60	57.1	899	2	Q8NIZ0	Q8niz0 neurospora
21	59.5	56.7	531	2	Q7PNQ9	Q7pnq9 anopheles g
22	59.5	56.7	647	2	Q89IE4	Q89ie4 bradyrhizob
23	59	56.2	347	2	Q9RKL9	Q9rkl9 streptomyce
24	59	56.2	389	2	Q9CM70	Q9cm70 pasteurella
25	59	56.2	1020	2	Q86PC3	Q86pc3 drosophila
26	59	56.2	1020	2	Q9W2J2	Q9w2j2 drosophila
27	59	56.2	1069	2	Q86BG1	Q86bg1 drosophila
28	58.5	55.7	181	2	Q64SR3	Q64sr3 bacteroides
29	58.5	55.7	496	2	Q8VQW6	Q8vqw6 azotobacter
30	58.5	55.7	508	2	Q9VGD2	Q9vgd2 drosophila
31	58.5	55.7	664	2	Q9VGD3	Q9vgd3 drosophila
32	58.5	55.7	694	2	Q8SWT7	Q8swt7 drosophila
33	58.5	55.7	1171	2	Q9P3E2	Q9p3e2 neurospora
34	58	55.2	92	2	Q9DF23	Q9df23 myoxocephal
35	58	55.2	124	2	Q7V6K8	Q7v6k8 prochloroco
36	58	55.2	387	2	O96113	O96113 plasmodium
37	58	55.2	558	2	Q7RWB7	Q7rwb7 neurospora
38	58	55.2	575	2	Q6PF71	Q6pf71 xenopus lae
39	58	55.2	660	2	Q88YM9	Q88ym9 lactobacill
40	58	55.2	809	2	P90534	P90534 dictyosteli
41	58	55.2	2197	1	POL1_TORVR	P29150 t rna1 poly
42	57.5	54.8	638	2	Q89IE3	Q89ie3 bradyrhizob
43	57	54.3	190	2	O15860	O15860 plasmodium
44	57	54.3	394	2	Q8X965	Q8x965 escherichia
45	57	54.3	508	2	Q875A8	Q875a8 podospora a

ALIGNMENTS

RESULT 1
Q8RXD0
ID Q8RXD0 PRELIMINARY; PRT; 485 AA.
AC Q8RXD0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Auxilin-like protein (At4g12780).
GN Name=At4g12780;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P.,
RA Dale J.M., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY081334; AAL91223.1; -.
DR EMBL; BT009679; AAP81797.1; -.
DR HSSP; Q27974; 1NZ6.
DR InterPro; IPR001623; DnaJ_N.
DR SMART; SM00271; DnaJ; 1.
SQ SEQUENCE 485 AA; 54793 MW; 1054D1021DB52AD5 CRC64;

Query Match 67.6%; Score 71; DB 2; Length 485;
Best Local Similarity 68.0%; Pred. No. 2.1;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| | :|||| |||| ||| ||:|
Db        184 AAAGARDKAAKAAAEAREKAEKAAA 208
```

RESULT 2
Q9SU08
ID Q9SU08 PRELIMINARY; PRT; 924 AA.
AC Q9SU08;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

```
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Auxilin-like protein.
GN   Name=T20K18.130; Synonyms=AT4g12780;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,
RA   Bancroft I., Mewes H.W., Mayer K.F.X., Schueller C.;
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,
RA   Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL049640; CAB40995.1; -.
DR   EMBL; AL161534; CAB78320.1; -.
DR   PIR; T06636; T06636.
DR   HSSP; Q27974; 1NZ6.
DR   InterPro; IPR001623; DnaJ_N.
DR   SMART; SM00271; DnaJ; 1.
SQ   SEQUENCE   924 AA;  102223 MW;  26E22C7C831EFF9B CRC64;

  Query Match             67.6%;  Score 71;  DB 2;  Length 924;
  Best Local Similarity   68.0%;  Pred. No. 3.5;
  Matches   17;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| | :|||| |||| ||| ||:|
Db        603 AAAGARDKAAKAAAEAREKAEKAAA 627


RESULT 3
Q7Q0M9
ID   Q7Q0M9      PRELIMINARY;      PRT;   278 AA.
AC   Q7Q0M9;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   AgCP8317 (Fragment).
GN   Name=agCG54338; ORFNames=ENSANGG00000011932;
OS   Anopheles gambiae str. PEST.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX   NCBI_TaxID=180454;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PEST;
RA   Anopheles Genome Sequencing Consortium;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the ribosomal protein L13P family.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAAB01008980; EAA14246.1; -.
DR   HSSP; O59300; 1J3A.
DR   GO; GO:0015934; C:large ribosomal subunit; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   GO; GO:0006412; P:protein biosynthesis; IEA.
DR   InterPro; IPR005822; Ribosomal_L13.
DR   InterPro; IPR005755; Ribosomal_L13e/a.
DR   Pfam; PF00572; Ribosomal_L13; 1.
DR   ProDom; PD001791; Ribosomal_L13; 1.
DR   TIGRFAMs; TIGR01077; L13_A_E; 1.
DR   PROSITE; PS00783; RIBOSOMAL_L13; 1.
KW   Ribonucleoprotein; Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   278 AA;  31601 MW;  D47C71B78A302495 CRC64;

  Query Match             63.8%;  Score 67;  DB 2;  Length 278;
  Best Local Similarity   65.2%;  Pred. No. 3.8;
  Matches   15;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|:|| | ||||| || || ||:
Db        243 AAKAARKEAKYAARAAAKAKKAT 265


RESULT 4
O69907
ID   O69907      PRELIMINARY;      PRT;   168 AA.
AC   O69907;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein SCO5619.
GN   ORFNames=SC2E1.36;
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
DR   EMBL; AL939124; CAA19411.1; -.
DR   PIR; T34804; T34804.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   168 AA;  17934 MW;  72063B195040BD6E CRC64;

  Query Match             61.9%;  Score 65;  DB 2;  Length 168;
  Best Local Similarity   62.5%;  Pred. No. 4.4;
  Matches   15;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:|:||||| : || ||:||| |:
Db        106 AEAKAAEKAERAAARAAKKAAPAP 129


RESULT 5
RL19_AGRT5
ID   RL19_AGRT5     STANDARD;      PRT;   179 AA.
AC   Q8UBZ5;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   50S ribosomal protein L19.
GN   Name=rplS; OrderedLocusNames=Atu2703, AGR_C_4900;
OS   Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX   NCBI_TaxID=176299;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA   Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA   Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA   Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA   Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA   Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA   Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA   Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA   Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA   Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA   Nester E.W.;
RT   "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT   C58.";
RL   Science 294:2317-2323(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA   Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA   Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA   Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA   Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA   Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA   Cielo C., Slater S.;
RT   "Genome sequence of the plant pathogen and biotechnology agent
RT   Agrobacterium tumefaciens C58.";
RL   Science 294:2323-2328(2001).
CC   -!- FUNCTION: This protein is located at the 30S-50S ribosomal subunit
CC       interface and may play a role in the structure and function of the
CC       aminoacyl-tRNA binding site (By similarity).
CC   -!- SIMILARITY: Belongs to the ribosomal protein L19P family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE009216; AAL43684.1; -.
DR   EMBL; AE008183; AAK88423.1; -.
DR   PIR; AF2908; AF2908.
DR   PIR; F97683; F97683.
DR   HAMAP; MF_00402; -; 1.
DR   InterPro; IPR001857; Ribosomal_L19.
DR   Pfam; PF01245; Ribosomal_L19; 1.
DR   PRINTS; PR00061; RIBOSOMALL19.
DR   ProDom; PD002979; Ribosomal_L19; 1.
DR   TIGRFAMs; TIGR01024; rplS_bact; 1.
DR   PROSITE; PS01015; RIBOSOMAL_L19; 1.
KW   Complete proteome; Ribosomal protein.
SQ   SEQUENCE   179 AA;  19474 MW;  F3256BA44A5AD2D1 CRC64;

  Query Match             60.0%;  Score 63;  DB 1;  Length 179;
  Best Local Similarity   69.2%;  Pred. No. 7.8;
  Matches   18;  Conservative    3;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAE--AAEKAAKAX 24
              |:| ||||||  |||  |||:||||:
Db        149 AQALAAEKAAAEAAEAKAAEEAAKAA 174


RESULT 6
Q83SA1
ID   Q83SA1      PRELIMINARY;      PRT;   413 AA.
AC   Q83SA1; Q7C2Q4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Membrane spanning protein, required for outer membrane integrity.
GN   Name=tolA; OrderedLocusNames=S0571, SF0558;
OS   Shigella flexneri.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=623;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=301 / Serotype 2a;
RX   MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA   Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA   Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA   Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA   Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA   Yu J.;
RT   "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT   through comparison with genomes of Escherichia coli K12 and O157.";
RL   Nucleic Acids Res. 30:4432-4441(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2457T;
RX   MEDLINE=22590274; PubMed=12704152;
RX   DOI=10.1128/IAI.71.5.2775-2786.2003;
RA   Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA   Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA   Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA   Schwartz D.C., Blattner F.R.;
RT   "Complete genome sequence and comparative genomics of Shigella
RT   flexneri serotype 2a strain 2457T.";
RL   Infect. Immun. 71:2775-2786(2003).
DR   EMBL; AE015086; AAN42202.1; -.
DR   EMBL; AE016979; AAP16075.1; -.
DR   HSSP; P19934; 1TOL.
DR   InterPro; IPR010528; TolA.
DR   Pfam; PF06519; TolA; 1.
KW   Complete proteome.
SQ   SEQUENCE   413 AA;  42355 MW;  93E10F2C5DE60DE8 CRC64;

  Query Match             60.0%;  Score 63;  DB 2;  Length 413;
  Best Local Similarity   60.0%;  Pred. No. 15;
  Matches   15;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|:|||:||| ||  |:| |:|:|
Db        143 ADAKAAEEAAKKAAADAKKKAEAEA 167


RESULT 7
TOLA_ECOLI
ID   TOLA_ECOLI     STANDARD;      PRT;   421 AA.
AC   P19934;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   25-JAN-2005 (Rel. 46, Last annotation update)
DE   TolA protein.
GN   Name=tolA; Synonyms=cim, excC, lky; OrderedLocusNames=b0739;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / JM105;
RX   MEDLINE=90078104; PubMed=2687247;
RA   Levengood S.K., Webster R.E.;
RT   "Nucleotide sequences of the tolA and tolB genes and localization of
RT   their products, components of a multistep translocation system in
RT   Escherichia coli.";
RL   J. Bacteriol. 171:6600-6609(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
```

```
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97061202; PubMed=8905232;
RA   Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA   Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA   Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA   Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA   Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA   Yano M., Horiuchi T.;
RT   "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 12.7-28.0 min region on the linkage map.";
RL   DNA Res. 3:137-155(1996).
RN   [4]
RP   DOMAINS.
RX   MEDLINE=91296736; PubMed=2068069;
RA   Levengood S.K., Beyer W.F. Jr., Webster R.E.;
RT   "TolA: a membrane protein involved in colicin uptake contains an
RT   extended helical region.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:5939-5943(1991).
RN   [5]
RP   INTERACTION WITH PORINS.
RX   MEDLINE=97133271; PubMed=8978668;
RA   Derouiche R., Gavioli M., Benedetti H., Prilipov A., Lazdunski C.,
RA   Lloubes R.;
RT   "TolA central domain interacts with Escherichia coli porins.";
RL   EMBO J. 15:6408-6415(1996).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 298-421.
RX   MEDLINE=99332679; PubMed=10404600; DOI=10.1016/S0969-2126(99)80092-6;
RA   Lubkowski J., Hennecke F., Plueckthun A., Wlodawer A.;
RT   "Filamentous phage infection: crystal structure of g3p in complex with
RT   its coreceptor, the C-terminal domain of TolA.";
RL   Structure 7:711-722(1999).
CC   -!- FUNCTION: Involved in the tonB-independent uptake of group A
CC       colicins (colicins A, E1, E2, E3, and K). Necessary for the
CC       colicins to reach their respective targets after initial binding
CC       to the bacteria. Also involved in the translocation of
CC       bacteriophage DNA.
CC   -!- SUBUNIT: Interacts, via domain II, with porins ompC, phoE and
CC       lamB.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M28232; AAA24683.1; -.
DR   EMBL; U00096; AAC73833.1; -.
DR   EMBL; D90713; BAA35405.1; -.
DR   PIR; JV0057; JV0057.
DR   PDB; 1TOL; X-ray; A=-.
DR   EchoBASE; EB1000; -.
DR   EcoGene; EG11007; tolA.
DR   InterPro; IPR010528; TolA.
DR   Pfam; PF06519; TolA; 1.
KW   3D-structure; Bacteriocin transport; Complete proteome;
KW   Inner membrane; Protein transport; Repeat; Transmembrane; Transport.
FT   DOMAIN        1     13       Cytoplasmic (Potential).
FT   TRANSMEM     14     34       Potential.
FT   DOMAIN       35    421       Periplasmic (Potential).
FT   DOMAIN       48    310       Domain II (alpha-helical).
FT   DOMAIN      311    421       Domain III (functional).
FT   DOMAIN      224    292       13 X tandem repeats of [EDA]-K(1,2)-
FT                                A(2,4).
FT   REPEAT      224    229       1.
FT   REPEAT      230    234       2.
FT   REPEAT      235    240       3.
FT   REPEAT      241    245       4.
FT   REPEAT      246    250       5.
FT   REPEAT      251    255       6.
FT   REPEAT      256    260       7.
FT   REPEAT      261    266       8.
FT   REPEAT      267    271       9.
FT   REPEAT      272    277       10.
FT   REPEAT      278    282       11.
FT   REPEAT      283    287       12.
FT   REPEAT      288    292       13.
FT   DISULFID    363    388
FT   HELIX       335    349
FT   TURN        350    351
FT   TURN        353    354
FT   HELIX       355    358
FT   TURN        359    360
FT   STRAND      363    369
FT   TURN        371    372
FT   STRAND      375    383
FT   HELIX       385    397
FT   HELIX       406    412
FT   TURN        413    414
FT   STRAND      416    421
SQ   SEQUENCE   421 AA;  43156 MW;  8B2F52B4B97C655E CRC64;

  Query Match             60.0%;  Score 63;  DB 1;  Length 421;
  Best Local Similarity   60.0%;  Pred. No. 16;
  Matches   15;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|:|||:||| ||  |:| |:|:|
Db        151 ADAKAAEEAAKKAAADAKKKAEAEA 175


RESULT 8
Q8FJT1
ID   Q8FJT1      PRELIMINARY;      PRT;   421 AA.
AC   Q8FJT1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   TolA protein.
GN   Name=tolA; OrderedLocusNames=c0818;
OS   Escherichia coli O6.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=217992;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX   MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA   Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA   Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA   Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA   Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT   "Extensive mosaic structure revealed by the complete genome sequence
RT   of uropathogenic Escherichia coli.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR   EMBL; AE016757; AAN79291.1; -.
DR   HSSP; P19934; 1TOL.
DR   InterPro; IPR010528; TolA.
DR   Pfam; PF06519; TolA; 1.
KW   Complete proteome.
SQ   SEQUENCE   421 AA;  43184 MW;  DB296626F056D385 CRC64;

  Query Match             60.0%;  Score 63;  DB 2;  Length 421;
  Best Local Similarity   60.0%;  Pred. No. 16;
  Matches   15;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|:|||:||| ||  |:| |:|:|
```

```
Db        151 ADAKAAEEAAKKAAADAKKKAEAEA 175


RESULT 9
Q6N8X8
ID   Q6N8X8      PRELIMINARY;      PRT;   441 AA.
AC   Q6N8X8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   OmpA/MotB domain, possible porin precursor.
GN   OrderedLocusNames=RPA1774;
OS   Rhodopseudomonas palustris.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Rhodopseudomonas.
OX   NCBI_TaxID=1076;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGA009 / ATCC BAA-98;
RX   PubMed=14704707; DOI=10.1038/nbt923;
RA   Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA   Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA   Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA   Harrison F.H., Gibson J., Harwood C.S.;
RT   "Complete genome sequence of the metabolically versatile
RT   photosynthetic bacterium Rhodopseudomonas palustris.";
RL   Nat. Biotechnol. 22:55-61(2004).
CC   -!- SIMILARITY: Belongs to the ompA family.
DR   EMBL; BX572598; CAE27215.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0009279; C:outer membrane (sensu Gram-negative Bacteria); IEA.
DR   GO; GO:0015288; F:porin activity; IEA.
DR   InterPro; IPR006664; Bac_OmpA.
DR   InterPro; IPR006665; OmpA/MotB.
DR   Pfam; PF00691; OmpA; 1.
DR   PRINTS; PR01021; OMPADOMAIN.
DR   ProDom; PD000930; OmpA/MotB; 1.
KW   Complete proteome; Porin; Signal.
FT   SIGNAL       1     27       Potential.
SQ   SEQUENCE   441 AA;  44811 MW;  F6BB66A4AFE183A0 CRC64;

  Query Match             58.1%;  Score 61;  DB 2;  Length 441;
  Best Local Similarity   62.5%;  Pred. No. 28;
  Matches   15;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |::|| | ||| ||||| | | |:
Db        237 ADSEAKEAAAKAAAEAAAKEAAAK 260


RESULT 10
Q8ZNE5
ID   Q8ZNE5      PRELIMINARY;      PRT;   593 AA.
AC   Q8ZNE5;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Putative von Willebrand factor, vWF type A domain.
GN   Name=yfbK; OrderedLocusNames=STM2315;
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2;
RX   MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
DR   EMBL; AE008803; AAL21216.1; -.
DR   Pfam; PF00092; VWA; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Complete proteome.
SQ   SEQUENCE   593 AA;  64640 MW;  595CA58158968357 CRC64;

  Query Match             58.1%;  Score 61;  DB 2;  Length 593;
  Best Local Similarity   62.5%;  Pred. No. 35;
  Matches   15;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :||||::||| ||| ||  | |:|
Db         57 QAEAAQQAAKKAAEYAEAKALADA 80


RESULT 11
Q7Z3F5
ID   Q7Z3F5      PRELIMINARY;      PRT;   711 AA.
AC   Q7Z3F5;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein DKFZp686F06102.
GN   Name=DKFZp686F06102;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Human fetal kidney;
RA   Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA   Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX537939; CAD97910.1; -.
DR   InterPro; IPR001451; Hexapep_transf.
DR   PROSITE; PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
KW   Hypothetical protein.
SQ   SEQUENCE   711 AA;  61765 MW;  95B624A99B4A989B CRC64;

  Query Match             57.6%;  Score 60.5;  DB 2;  Length 711;
  Best Local Similarity   58.6%;  Pred. No. 47;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db        446 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 474


RESULT 12
ELS_HUMAN
ID   ELS_HUMAN      STANDARD;      PRT;   730 AA.
AC   P15502; Q14233; Q14238;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Elastin precursor (Tropoelastin).
GN   Name=ELN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM B).
RX   MEDLINE=87289668; PubMed=3039501;
RA   Indik Z., Yeh H., Ornstein-Goldstein N., Sheppard P., Anderson N.,
RA   Rosenbloom J.C., Peltonen L., Rosenbloom J.;
RT   "Alternative splicing of human elastin mRNA indicated by sequence
RT   analysis of cloned genomic and complementary DNA.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:5680-5684(1987).
```

```
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Skin fibroblast;
RX   MEDLINE=89009960; PubMed=3171221;
RA   Fazio M.J., Olsen D.R., Kauh E.A., Baldwin C.T., Indik Z.,
RA   Ornstein-Goldstein N., Yeh H., Rosenbloom J., Uitto J.;
RT   "Cloning of full-length elastin cDNAs from a human skin fibroblast
RT   recombinant cDNA library: further elucidation of alternative splicing
RT   utilizing exon-specific oligonucleotides.";
RL   J. Invest. Dermatol. 91:458-464(1988).
RN   [3]
RP   SEQUENCE OF 164-724 FROM N.A. (ISOFORM B).
RC   TISSUE=Placenta;
RX   MEDLINE=88156138; PubMed=2831431;
RA   Fazio M.J., Olsen D.R., Kuivaniemi H., Chu M.L., Davidson J.M.,
RA   Rosenbloom J., Uitto J.;
RT   "Isolation and characterization of human elastin cDNAs, and age-
RT   associated variation in elastin gene expression in cultured skin
RT   fibroblasts.";
RL   Lab. Invest. 58:270-277(1988).
RN   [4]
RP   SEQUENCE OF 603-730 FROM N.A.
RC   TISSUE=Hippocampus, and Placenta;
RX   MEDLINE=96291399; PubMed=8689688; DOI=10.1016/S0092-8674(00)80077-X;
RA   Frangiskakis J.M., Ewart A.K., Morris C.A., Mervis C.B., Bertrand J.,
RA   Robinson B.F., Klein B.P., Ensing G.J., Everett L.A., Green E.D.,
RA   Proeschel C., Gutowski N.J., Noble M., Atkinson D.L., Odelberg S.J.,
RA   Keating M.T.;
RT   "LIM-kinase1 hemizygosity implicated in impaired visuospatial
RT   constructive cognition.";
RL   Cell 86:59-69(1996).
RN   [5]
RP   INVOLVEMENT IN CUTIS LAXA.
RX   MEDLINE=99091639; PubMed=9873040; DOI=10.1074/jbc.274.2.981;
RA   Zhang M.-C., He L., Giro M., Yong S.L., Tiller G.E., Davidson J.M.;
RT   "Cutis laxa arising from frameshift mutations in exon 30 of the
RT   elastin gene (ELN).";
RL   J. Biol. Chem. 274:981-986(1999).
RN   [6]
RP   INVOLVEMENT IN SVAS.
RX   PubMed=10942104;
RA   Urban Z., Michels V.V., Thibodeau S.N., Davis E.C., Bonnefont J.-P.,
RA   Munnich A., Eyskens B., Gewillig M., Devriendt K., Boyd C.D.;
RT   "Isolated supravalvular aortic stenosis: functional haploinsufficiency
RT   of the elastin gene as a result of nonsense-mediated decay.";
RL   Hum. Genet. 106:577-588(2000).
CC   -!- FUNCTION: Major structural protein of tissues such as aorta and
CC       nuchal ligament, which must expand rapidly and recover completely.
CC   -!- SUBUNIT: The polymeric elastin chains are cross-linked together
CC       into an extensible 3D network.
CC   -!- SUBCELLULAR LOCATION: Extracellular matrix of elastic fibers.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms seem to exist;
CC       Name=1;
CC         IsoId=P15502-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P15502-2; Sequence=VSP_004243;
CC   -!- PTM: The crosslinks are made of deaminated Lys.
CC   -!- DISEASE: Defects in ELN are a cause of autosomal dominant cutis
CC       laxa [MIM:123700]. Cutis laxa is a rare connective tissue disorder
CC       characterized by loose, hyperextensible skin with decreased
CC       resilience and elasticity leading to a premature aged appearance.
CC       The skin changes are often accompanied by extracutaneous
CC       manifestations, including pulmonary emphysema, bladder
CC       diverticula, pulmonary artery stenosis and pyloric stenosis.
CC   -!- DISEASE: Haploinsufficiency of ELN may be the cause of certain
CC       cardiovascular and musculo-skeletal abnormalities observed in
CC       Williams-Beuren syndrome (WBS) [MIM:194050]. WBS is a rare
CC       developmental disorder and a contiguous gene deletion syndrome
CC       involving genes from chromosome band 7q11.23.
CC   -!- DISEASE: Defects in ELN are the cause of supravalvular aortic
CC       stenosis (SVAS) [MIM:185500]. SVAS is a congenital narrowing of
CC       the ascending aorta which can occur sporadically, as an autosomal
CC       dominant condition, or as one component of Williams-Beuren
CC       syndrome.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M17282; AAC98394.1; -.
DR   EMBL; M16983; AAC98394.1; JOINED.
DR   EMBL; M17265; AAC98394.1; JOINED.
DR   EMBL; M17266; AAC98394.1; JOINED.
DR   EMBL; M17267; AAC98394.1; JOINED.
DR   EMBL; M17268; AAC98394.1; JOINED.
DR   EMBL; M17270; AAC98394.1; JOINED.
DR   EMBL; M17271; AAC98394.1; JOINED.
DR   EMBL; M17272; AAC98394.1; JOINED.
DR   EMBL; M17273; AAC98394.1; JOINED.
DR   EMBL; M17275; AAC98394.1; JOINED.
DR   EMBL; M17276; AAC98394.1; JOINED.
DR   EMBL; M17277; AAC98394.1; JOINED.
DR   EMBL; M17278; AAC98394.1; JOINED.
DR   EMBL; M17279; AAC98394.1; JOINED.
DR   EMBL; M17280; AAC98394.1; JOINED.
DR   EMBL; M17281; AAC98394.1; JOINED.
DR   EMBL; M36860; AAA52382.1; -.
DR   EMBL; M24782; AAA53190.1; -.
DR   EMBL; U62292; AAB17544.1; -.
DR   EMBL; X15603; CAA33627.1; -.
DR   PIR; A32707; EAHU.
DR   HSSP; P50099; 1ZFJ.
DR   Genew; HGNC:3327; ELN.
DR   MIM; 130160; -.
DR   MIM; 123700; -.
DR   MIM; 194050; -.
DR   MIM; 185500; -.
DR   GO; GO:0005578; C:extracellular matrix; TAS.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0005201; F:extracellular matrix structural constituent; TAS.
DR   GO; GO:0008283; P:cell proliferation; TAS.
DR   GO; GO:0008015; P:circulation; TAS.
DR   GO; GO:0009887; P:organogenesis; TAS.
DR   GO; GO:0007585; P:respiratory gaseous exchange; TAS.
DR   InterPro; IPR003979; tropoelastin.
DR   PRINTS; PR01500; TROPOELASTIN.
KW   Alternative splicing; Repeat; Signal; Structural protein;
KW   Williams-Beuren syndrome.
FT   SIGNAL        1     26
FT   CHAIN        27    730       Elastin.
FT   DISULFID    720    725       By similarity.
FT   VARSPLIC    472    477       Missing (in isoform 2).
FT                                /FTId=VSP_004243.
SQ   SEQUENCE   730 AA;  63260 MW;  AB06D15BA567AE46 CRC64;

  Query Match             57.6%;  Score 60.5;  DB 1;  Length 730;
  Best Local Similarity   58.6%;  Pred. No. 48;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db        441 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 469


RESULT 13
Q14234
ID   Q14234      PRELIMINARY;      PRT;   757 AA.
AC   Q14234;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
```

```
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Elastin.
GN   Name=ELN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87289668; PubMed=3039501;
RA   Indik Z., Yeh H., Ornstein-Goldstein N., Sheppard P., Anderson N.,
RA   Rosenbloom J.C., Peltonen L., Rosenbloom J.;
RT   "Alternative splicing of human elastin mRNA indicated by sequence
RT   analysis of cloned genomic and complementary DNA.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:5680-5684(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87274906; PubMed=3038460;
RA   Indik Z., Yoon K., Morrow S.D., Cicila G., Rosenbloom J.,
RA   Rosenbloom J., Ornstein-Goldstein N.;
RT   "Structure of the 3' region of the human elastin gene: great abundance
RT   of Alu repetitive sequences and few coding sequences.";
RL   Connect. Tissue Res. 16:197-211(1987).
DR   EMBL; M17282; AAC98395.1; -.
DR   EMBL; M16983; AAC98395.1; JOINED.
DR   EMBL; M17265; AAC98395.1; JOINED.
DR   EMBL; M17266; AAC98395.1; JOINED.
DR   EMBL; M17267; AAC98395.1; JOINED.
DR   EMBL; M17268; AAC98395.1; JOINED.
DR   EMBL; M17270; AAC98395.1; JOINED.
DR   EMBL; M17271; AAC98395.1; JOINED.
DR   EMBL; M17272; AAC98395.1; JOINED.
DR   EMBL; M17273; AAC98395.1; JOINED.
DR   EMBL; M17274; AAC98395.1; JOINED.
DR   EMBL; M17275; AAC98395.1; JOINED.
DR   EMBL; M17276; AAC98395.1; JOINED.
DR   EMBL; M17277; AAC98395.1; JOINED.
DR   EMBL; M17278; AAC98395.1; JOINED.
DR   EMBL; M17279; AAC98395.1; JOINED.
DR   EMBL; M17280; AAC98395.1; JOINED.
DR   EMBL; M17281; AAC98395.1; JOINED.
DR   GO; GO:0005578; C:extracellular matrix (sensu Metazoa); NAS.
DR   GO; GO:0030023; F:extracellular matrix constituent conferring. . .; NAS.
DR   InterPro; IPR001451; Hexapep_transf.
DR   InterPro; IPR003979; tropoelastin.
DR   PRINTS; PR01500; TROPOELASTIN.
DR   PROSITE; PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
SQ   SEQUENCE   757 AA;  66136 MW;  23B7FE5B8AF85CA8 CRC64;

  Query Match             57.6%;  Score 60.5;  DB 2;  Length 757;
  Best Local Similarity   58.6%;  Pred. No. 49;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db        441 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 469


RESULT 14
Q75MU5
ID   Q75MU5      PRELIMINARY;      PRT;   757 AA.
AC   Q75MU5;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein ELN.
GN   Name=ELN;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;
RA   Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
RA   Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
RA   Wylie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E.,
RA   Fewell G.A., Delehaunty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
RA   Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,
RA   Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,
RA   Ozersky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
RA   Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
RA   Kozlowicz-Reilly A., Leonard S., Rohlfing T., Rock S.M.,
RA   Tin-Wollam A.M., Abbott A., Minx P., Maupin R., Strowmatt C.,
RA   Latreille P., Miller N., Johnson D., Murray J., Woessner J.P.,
RA   Wendl M.C., Yang S.P., Schultz B.R., Wallis J.W., Spieth J.,
RA   Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.L.,
RA   Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,
RA   Clifton S.W., Chissoe S.L., Marra M.A., Raymond C., Haugen E.,
RA   Gillett W., Zhou Y., James R., Phelps K., Iadanoto S., Bubb K.,
RA   Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Furey T.S.,
RA   Baertsch R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,
RA   Bailey J.A., Portnoy M.E., Torrents D., Chinwalla A.T., Gish W.R.,
RA   Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
RA   Waterston R.H., Wilson R.K.;
RT   "The DNA sequence of human chromosome 7.";
RL   Nature 424:157-164(2003).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Waterston R.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Wilson R.;
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC005056; AAS07435.1; -.
DR   GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.
DR   GO; GO:0005201; F:extracellular matrix structural constituent; IEA.
DR   InterPro; IPR001451; Hexapep_transf.
DR   InterPro; IPR003979; tropoelastin.
DR   PRINTS; PR01500; TROPOELASTIN.
DR   PROSITE; PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
KW   Hypothetical protein.
SQ   SEQUENCE   757 AA;  66106 MW;  2B24F955D8360738 CRC64;

  Query Match             57.6%;  Score 60.5;  DB 2;  Length 757;
  Best Local Similarity   58.6%;  Pred. No. 49;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db        441 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 469


RESULT 15
Q6N503
ID   Q6N503      PRELIMINARY;      PRT;   105 AA.
AC   Q6N503;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=RPA3180;
OS   Rhodopseudomonas palustris.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Rhodopseudomonas.
OX   NCBI_TaxID=1076;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGA009 / ATCC BAA-98;
RX   PubMed=14704707; DOI=10.1038/nbt923;
RA   Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA   Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA   Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA   Harrison F.H., Gibson J., Harwood C.S.;
RT   "Complete genome sequence of the metabolically versatile
```

```
RT   photosynthetic bacterium Rhodopseudomonas palustris.";
RL   Nat. Biotechnol. 22:55-61(2004).
DR   EMBL; BX572603; CAE28621.1; -.
DR   GO; GO:0000786; C:nucleosome; IEA.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006334; P:nucleosome assembly; IEA.
DR   InterPro; IPR005819; Histone_H5.
DR   PRINTS; PR00624; HISTONEH5.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   105 AA;  11042 MW;  CEDB59B3D937E980 CRC64;

  Query Match             57.1%;  Score 60;  DB 2;  Length 105;
  Best Local Similarity   56.0%;  Pred. No. 11;
  Matches   14;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |: :||:|||| | : |:|||| :|
Db         58 AAKKAAKKAAKKAPKKAKKAAKKAA 82


RESULT 16
Q9AB65
ID   Q9AB65      PRELIMINARY;      PRT;   177 AA.
AC   Q9AB65;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   ATP synthase F0, B' subunit.
GN   OrderedLocusNames=CC0366;
OS   Caulobacter crescentus.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC   Caulobacteraceae; Caulobacter.
OX   NCBI_TaxID=155892;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 19089 / CB15;
RX   MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA   Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA   Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA   Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA   Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA   Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA   Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA   Fraser C.M.;
RT   "Complete genome sequence of Caulobacter crescentus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC   -!- SIMILARITY: Belongs to the ATPase B chain family.
DR   EMBL; AE005710; AAK22353.1; -.
DR   PIR; E87294; E87294.
DR   TIGR; CC0366; -.
DR   GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
DR   GO; GO:0015078; F:hydrogen ion transporter activity; IEA.
DR   GO; GO:0016820; F:hydrolase activity, acting on acid anhydrid. . .; IEA.
DR   GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.
DR   GO; GO:0015992; P:proton transport; IEA.
DR   InterPro; IPR002146; ATPsynt_B/B'sub.
DR   Pfam; PF00430; ATP-synt_B; 1.
KW   CF(0); Complete proteome; Hydrogen ion transport; Ion transport;
KW   Transmembrane; Transport.
SQ   SEQUENCE   177 AA;  18465 MW;  6F0A2E32CC3D2912 CRC64;

  Query Match             57.1%;  Score 60;  DB 2;  Length 177;
  Best Local Similarity   60.0%;  Pred. No. 17;
  Matches   15;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|||||: ||  |  ||| | |:|
Db        110 ASAEAAERQAKEEAVLAEKLAAAEA 134


RESULT 17
Q6N4V4
ID   Q6N4V4      PRELIMINARY;      PRT;   371 AA.
AC   Q6N4V4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Adenylate kinase (EC 2.7.4.3).
GN   Name=adk; OrderedLocusNames=RPA3229;
OS   Rhodopseudomonas palustris.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Rhodopseudomonas.
OX   NCBI_TaxID=1076;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGA009 / ATCC BAA-98;
RX   PubMed=14704707; DOI=10.1038/nbt923;
RA   Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA   Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA   Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA   Harrison F.H., Gibson J., Harwood C.S.;
RT   "Complete genome sequence of the metabolically versatile
RT   photosynthetic bacterium Rhodopseudomonas palustris.";
RL   Nat. Biotechnol. 22:55-61(2004).
CC   -!- FUNCTION: This small ubiquitous enzyme is essential for
CC       maintenance and cell growth (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + AMP = 2 ADP.
CC   -!- SUBUNIT: Monomer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the adenylate kinase family.
DR   EMBL; BX572603; CAE28670.1; -.
DR   HSSP; P05082; 1AKE.
DR   GO; GO:0004017; F:adenylate kinase activity; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0016301; F:kinase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   InterPro; IPR000850; Adenylate_kin.
DR   InterPro; IPR006259; Adenyl_kin_sub.
DR   Pfam; PF00406; ADK; 1.
DR   PRINTS; PR00094; ADENYLTKNASE.
DR   ProDom; PD000657; Adenylate_kin; 1.
DR   TIGRFAMs; TIGR01351; adk; 1.
DR   PROSITE; PS00113; ADENYLATE_KINASE; 1.
KW   ATP-binding; Complete proteome; Kinase; Transferase.
SQ   SEQUENCE   371 AA;  37905 MW;  9BB86A147A346206 CRC64;

  Query Match             57.1%;  Score 60;  DB 2;  Length 371;
  Best Local Similarity   56.0%;  Pred. No. 32;
  Matches   14;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|:| :|||| | :||: ||| :|
Db        306 AVAKAGKKAAKAAVKAAKGAAKPTA 330


RESULT 18
Q6NH65
ID   Q6NH65      PRELIMINARY;      PRT;   572 AA.
AC   Q6NH65;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Putative invasion protein.
GN   OrderedLocusNames=DIP1281;
OS   Corynebacterium diphtheriae.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=1717;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Biotype gravis / NCTC 13129;
RX   MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;
RA   Cerdeno-Tarraga A.-M., Efstratiou A., Dover L.G., Holden M.T.G.,
RA   Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,
```

```
RA   De Zoysa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA   Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
RA   Rabbinowitsch E., Rutherford K.M., Thomson N.R., Unwin L.,
RA   Whitehead S., Barrell B.G., Parkhill J.;
RT   "The complete genome sequence and analysis of Corynebacterium
RT   diphtheriae NCTC13129.";
RL   Nucleic Acids Res. 31:6516-6523(2003).
DR   EMBL; BX248357; CAE49808.1; -.
DR   InterPro; IPR000064; NLP_P60.
DR   Pfam; PF00877; NLPC_P60; 1.
KW   Complete proteome.
SQ   SEQUENCE   572 AA;  61149 MW;  6ECCE6CD145263F2 CRC64;

  Query Match             57.1%;  Score 60;  DB 2;  Length 572;
  Best Local Similarity   56.0%;  Pred. No. 45;
  Matches   14;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |: :| |:||: ||||| |||  :|
Db        294 AARKAKEEAARIAAEAARKAADEAA 318


RESULT 19
Q6UBQ3
ID   Q6UBQ3      PRELIMINARY;      PRT;   738 AA.
AC   Q6UBQ3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Radial spoke protein 2.
OS   Chlamydomonas reinhardtii.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC   Chlamydomonadaceae; Chlamydomonas.
OX   NCBI_TaxID=3055;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=14871938; DOI=10.1128/EC.3.1.72-81.2004;
RA   Yang P., Yang C., Sale W.S.;
RT   "Flagellar radial spoke protein 2 is a calmodulin binding protein
RT   required for motility in Chlamydomonas reinhardtii.";
RL   Eukaryotic Cell 3:72-81(2004).
DR   EMBL; AY373262; AAQ92371.1; -.
DR   InterPro; IPR007858; Dpy-30.
DR   Pfam; PF05186; Dpy-30; 1.
SQ   SEQUENCE   738 AA;  77362 MW;  506811B4975539AD CRC64;

  Query Match             57.1%;  Score 60;  DB 2;  Length 738;
  Best Local Similarity   62.5%;  Pred. No. 55;
  Matches   15;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||: |||||  ||:|:|
Db        651 AAAAAEAAAEAAAEAAAAAAEAAA 674


RESULT 20
Q8NIZ0
ID   Q8NIZ0      PRELIMINARY;      PRT;   899 AA.
AC   Q8NIZ0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Related to kinetoplast-associated protein KAP.
GN   Name=5F3.190;
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA   Nyakatura G., Mewes H.W., Mannhaupt G.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   German Neurospora genome project;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL807371; CAD37020.1; -.
DR   HSSP; P42639; 1C1G.
SQ   SEQUENCE   899 AA;  99309 MW;  5A110FCA4C09D8F9 CRC64;

  Query Match             57.1%;  Score 60;  DB 2;  Length 899;
  Best Local Similarity   73.7%;  Pred. No. 65;
  Matches   14;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 EAAEKAAKYAAEAAEKAAK 22
              :|||:||| ||| |||| |
Db        511 KAAEEAAKKAAEEAEKARK 529


RESULT 21
Q7PNQ9
ID   Q7PNQ9      PRELIMINARY;      PRT;   531 AA.
AC   Q7PNQ9;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   ENSANGP00000017612 (Fragment).
GN   Name=ENSANGG00000015123;
OS   Anopheles gambiae str. PEST.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX   NCBI_TaxID=180454;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PEST;
RA   Anopheles Genome Sequencing Consortium;
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAAB01008960; EAA11760.2; -.
DR   HSSP; P04802; 1ASZ.
DR   GO; GO:0005737; C:cytoplasm; IEA.
DR   GO; GO:0004815; F:aspartate-tRNA ligase activity; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0006422; P:aspartyl-tRNA aminoacylation; IEA.
DR   InterPro; IPR004523; AspS_arch.
DR   InterPro; IPR008994; Nucleic_acid_OB.
DR   InterPro; IPR004364; tRNA-synt_2.
DR   InterPro; IPR002312; tRNA-synt_asp.
DR   InterPro; IPR004365; tRNA_anti.
DR   InterPro; IPR006195; tRNA_ligase_II.
DR   Pfam; PF00152; tRNA-synt_2; 1.
DR   Pfam; PF01336; tRNA_anti; 1.
DR   PRINTS; PR01042; TRNASYNTHASP.
DR   TIGRFAMs; TIGR00458; aspS_arch; 1.
DR   PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   531 AA;  58958 MW;  A9468D3F53448317 CRC64;

  Query Match             56.7%;  Score 59.5;  DB 2;  Length 531;
  Best Local Similarity   69.6%;  Pred. No. 48;
  Matches   16;  Conservative    4;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 AXAEA-AEKAAKYAAEAAEKAAK 22
              |:||| ::|||| ||: ||||||
Db          1 AGAEATSKKAAKKAAKDAEKAAK 23


RESULT 22
Q89IE4
ID   Q89IE4      PRELIMINARY;      PRT;   647 AA.
AC   Q89IE4;
```

```
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Bll5695 protein.
GN   OrderedLocusNames=bll5695;
OS   Bradyrhizobium japonicum.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Bradyrhizobium.
OX   NCBI_TaxID=375;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=USDA110;
RX   MEDLINE=22484998; PubMed=12597275;
RA   Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA   Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA   Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA   Tabata S.;
RT   "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT   Bradyrhizobium japonicum USDA110.";
RL   DNA Res. 9:189-197(2002).
DR   EMBL; AP005955; BAC50960.1; -.
DR   GO; GO:0030693; F:caspase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001309; ICE_p20.
DR   PROSITE; PS50208; CASPASE_P20; 1.
KW   Complete proteome.
SQ   SEQUENCE   647 AA;  69607 MW;  69FB4EDECF6CF836 CRC64;

  Query Match             56.7%;  Score 59.5;  DB 2;  Length 647;
  Best Local Similarity   53.3%;  Pred. No. 57;
  Matches   16;  Conservative    5;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEK-----AAKYAAEAAEKAAKAXA 25
              |: :||||     ||| ||| ||:||: :|
Db        425 AEKQAAEKVKAELAAKQAAEKAEQAARPAA 454


RESULT 23
Q9RKL9
ID   Q9RKL9      PRELIMINARY;      PRT;   347 AA.
AC   Q9RKL9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Probable peptidase.
GN   ORFNames=SCD17.12;
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
DR   EMBL; AL939118; CAB56389.1; -.
DR   GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR002886; Peptidase_M23B.
DR   InterPro; IPR011054; Rudmnt_hyb_motif.
DR   Pfam; PF01551; Peptidase_M23; 1.
KW   Complete proteome.
SQ   SEQUENCE   347 AA;  35432 MW;  456DFC61B6C2FF0D CRC64;

  Query Match             56.2%;  Score 59;  DB 2;  Length 347;
  Best Local Similarity   68.2%;  Pred. No. 39;
  Matches   15;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAK 22
              |:||||  | | ||||| || |
Db        163 AAAEAAAAAAEKKAAEAAAKAEK 184


RESULT 24
Q9CM70
ID   Q9CM70      PRELIMINARY;      PRT;   389 AA.
AC   Q9CM70;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   TolA.
GN   Name=tolA; OrderedLocusNames=PM0968;
OS   Pasteurella multocida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Pasteurella.
OX   NCBI_TaxID=747;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Pm70;
RX   MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA   May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT   "Complete genomic sequence of Pasteurella multocida Pm70.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR   EMBL; AE006136; AAK03052.1; -.
DR   HSSP; P01096; 1GMJ.
DR   InterPro; IPR009148; SibA.
DR   InterPro; IPR010528; TolA.
DR   InterPro; IPR000533; Tropomyosin.
DR   Pfam; PF06519; TolA; 1.
DR   PRINTS; PR01852; SIBAPROTEIN.
DR   PRINTS; PR00194; TROPOMYOSIN.
KW   Complete proteome.
SQ   SEQUENCE   389 AA;  42197 MW;  B4032F2A2FD9E94B CRC64;

  Query Match             56.2%;  Score 59;  DB 2;  Length 389;
  Best Local Similarity   60.0%;  Pred. No. 43;
  Matches   15;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|||  || |  |||  || ||:|
Db        224 AEAEAKAKAEKAKAEAKAKAEKAKA 248


RESULT 25
Q86PC3
ID   Q86PC3      PRELIMINARY;      PRT;  1020 AA.
AC   Q86PC3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   RE13301p.
GN   ORFNames=CG18375;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
```

```
RA   Celniker S.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 SH3 domain.
DR   EMBL; BT003215; AAO24970.1; -.
DR   HSSP; P06241; 1SHF.
DR   FlyBase; FBgn0034606; CG18375.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001452; SH3.
DR   Pfam; PF00023; Ank; 2.
DR   Pfam; PF00018; SH3_1; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   ProDom; PD000066; SH3; 1.
DR   SMART; SM00248; ANK; 2.
DR   SMART; SM00326; SH3; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 2.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50002; SH3; 1.
KW   ANK repeat; SH3 domain.
SQ   SEQUENCE   1020 AA;  110433 MW;  42A3AE30EC71787B CRC64;

  Query Match             56.2%;  Score 59;  DB 2;  Length 1020;
  Best Local Similarity   60.0%;  Pred. No. 93;
  Matches   15;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
                |:| ||  ||: ||||| :|| |:|
Db        462 AAAAAAAAAAQAAAEAANQAATAAA 486


RESULT 26
Q9W2J2
ID   Q9W2J2      PRELIMINARY;      PRT;  1020 AA.
AC   Q9W2J2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   CG18375-PA.
GN   ORFNames=CG18375;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfied E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN   [5]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 SH3 domain.
DR   EMBL; AE003453; AAF46699.3; -.
DR   HSSP; P06241; 1SHF.
DR   IntAct; Q9W2J2; -.
DR   FlyBase; FBgn0034606; CG18375.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001452; SH3.
DR   Pfam; PF00023; Ank; 2.
DR   PRINTS; PR01415; ANKYRIN.
DR   ProDom; PD000066; SH3; 1.
DR   SMART; SM00248; ANK; 2.
DR   SMART; SM00326; SH3; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 2.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50002; SH3; 1.
KW   ANK repeat; SH3 domain.
SQ   SEQUENCE   1020 AA;  110374 MW;  B18C928C514313DC CRC64;

  Query Match             56.2%;  Score 59;  DB 2;  Length 1020;
  Best Local Similarity   60.0%;  Pred. No. 93;
```

```
  Matches   15;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||  ||: ||||| :|| |:|
Db        462 AAAAAAAAAAQAAAEAANQAATAAA 486


RESULT 27
Q86BG1
ID   Q86BG1      PRELIMINARY;      PRT;  1069 AA.
AC   Q86BG1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   CG18375-PB.
GN   ORFNames=CG18375;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfied E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN   [5]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 SH3 domain.
DR   EMBL; AE003453; AAO41341.1; -.
DR   HSSP; P06241; 1SHF.
DR   FlyBase; FBgn0034606; CG18375.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001452; SH3.
DR   Pfam; PF00023; Ank; 2.
DR   Pfam; PF00018; SH3_1; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   ProDom; PD000066; SH3; 1.
DR   SMART; SM00248; ANK; 2.
DR   SMART; SM00326; SH3; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 2.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50002; SH3; 1.
KW   ANK repeat; SH3 domain.
SQ   SEQUENCE   1069 AA;  115158 MW;  BF102B0C044F80DA CRC64;

  Query Match             56.2%;  Score 59;  DB 2;  Length 1069;
  Best Local Similarity   60.0%;  Pred. No. 97;
  Matches   15;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||  ||: ||||| :|| |:|
Db        511 AAAAAAAAAAQAAAEAANQAATAAA 535


RESULT 28
Q64SR3
ID   Q64SR3      PRELIMINARY;      PRT;   181 AA.
AC   Q64SR3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   30S ribosomal protein S16.
GN   ORFNames=BF2716;
OS   Bacteroides fragilis.
OC   Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC   Bacteroidaceae; Bacteroides.
OX   NCBI_TaxID=817;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=YCH46;
RA   Kuwahara T., Yamashita A., Hirakawa H., Nakayama H., Toh H., Okada N.,
RA   Kuhara S., Hattori M., Hayashi T., Ohnishi Y.;
RT   "Genomic analysis of Bacteroides fragilis reveals extensive DNA
RT   inversions regulating cell surface adaptation.";
RL   Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2004).
DR   EMBL; AP006841; BAD49466.1; -.
KW   Ribosomal protein.
SQ   SEQUENCE   181 AA;  19609 MW;  35B33BF4EC4DCAD3 CRC64;

  Query Match             55.7%;  Score 58.5;  DB 2;  Length 181;
  Best Local Similarity   64.0%;  Pred. No. 26;
  Matches   16;  Conservative    5;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEA-AEKAAKAX 24
              |: :|||:||| |||| ||:|| |:
Db        148 AEKKAAEEAAKAAAEAPAEEAAPAE 172


RESULT 29
Q8VQW6
ID   Q8VQW6      PRELIMINARY;      PRT;   496 AA.
AC   Q8VQW6;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   RnfC.
GN   Name=rnfC;
OS   Azotobacter vinelandii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Azotobacter.
OX   NCBI_TaxID=354;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rubio L.M., Brown C.S., Ludden P.W.;
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF450501; AAL47174.1; -.
DR   HSSP; Q45560; 1BWE.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005489; F:electron transporter activity; IEA.
DR   GO; GO:0005506; F:iron ion binding; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   Pfam; PF01512; Complex1_51K; 1.
DR   Pfam; PF00037; Fer4; 2.
DR   TIGRFAMs; TIGR01945; rnfC; 1.
DR   PROSITE; PS00198; 4FE4S_FERREDOXIN; 2.
KW   4Fe-4S; Iron; Iron-sulfur; Metal-binding.
SQ   SEQUENCE   496 AA;  52171 MW;  0F153E1B83A10E5B CRC64;

  Query Match             55.7%;  Score 58.5;  DB 2;  Length 496;
  Best Local Similarity   66.7%;  Pred. No. 60;
  Matches   16;  Conservative    4;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :|:|| |||| ||:|| |||| :|
Db        465 AAKAAAKAAKEAAKAA-KAAKTKA 487


RESULT 30
Q9VGD2
ID   Q9VGD2      PRELIMINARY;      PRT;   508 AA.
AC   Q9VGD2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   CG31361-PB.
GN   ORFNames=CG31361;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfied E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
```

```
RL   Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN   [5]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003694; AAF54751.2; -.
DR   FlyBase; FBgn0051361; CG31361.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00408; IGc2; 1.
DR   PROSITE; PS50835; IG_LIKE; 2.
SQ   SEQUENCE   508 AA;  54018 MW;  203110CA2A9523EE CRC64;

  Query Match             55.7%;  Score 58.5;  DB 2;  Length 508;
  Best Local Similarity   66.7%;  Pred. No. 61;
  Matches   16;  Conservative    4;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:|:||| ||| ||||| :|| |:
Db        187 AAADAAE-AAKLAAEAAAQAAAAK 209


RESULT 31
Q9VGD3
ID   Q9VGD3      PRELIMINARY;      PRT;   664 AA.
AC   Q9VGD3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   CG31361-PA.
GN   ORFNames=CG31361;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfied E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN   [5]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RG   FlyBase;
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003694; AAF54750.2; -.
DR   IntAct; Q9VGD3; -.
DR   FlyBase; FBgn0051361; CG31361.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00408; IGc2; 1.
DR   PROSITE; PS50835; IG_LIKE; 2.
SQ   SEQUENCE   664 AA;  70818 MW;  75C8A45055CF457B CRC64;

  Query Match             55.7%;  Score 58.5;  DB 2;  Length 664;
  Best Local Similarity   66.7%;  Pred. No. 75;
  Matches   16;  Conservative    4;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:|:||| ||| ||||| :|| |:
Db        343 AAADAAE-AAKLAAEAAAQAAAAK 365


RESULT 32
Q8SWT7
ID   Q8SWT7      PRELIMINARY;      PRT;   694 AA.
```

```
AC   Q8SWT7;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   RE56367p.
GN   Name=CG14738; ORFNames=CG31361;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY095087; AAM11415.1; -.
DR   HSSP; Q9UQH9; 1E0O.
DR   FlyBase; FBgn0051361; CG31361.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00408; IGc2; 1.
DR   PROSITE; PS50835; IG_LIKE; 2.
SQ   SEQUENCE   694 AA;  75164 MW;  9C242F9F03051491 CRC64;

  Query Match             55.7%;  Score 58.5;  DB 2;  Length 694;
  Best Local Similarity   66.7%;  Pred. No. 78;
  Matches   16;  Conservative    4;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:|:||| ||| ||||| :|| |:
Db        343 AAADAAE-AAKLAAEAAAQAAAAK 365


RESULT 33
Q9P3E2
ID   Q9P3E2      PRELIMINARY;      PRT;  1171 AA.
AC   Q9P3E2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Related to transport protein USO1.
GN   Name=B13I18.10;
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA   Nyakatura G., Mewes H.W., Mannhaupt G.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   German Neurospora genome project;
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL390189; CAB99171.1; -.
DR   GO; GO:0005737; C:cytoplasm; IEA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008565; F:protein transporter activity; IEA.
DR   GO; GO:0006886; P:intracellular protein transport; IEA.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR006955; Uso1_p115_C.
DR   InterPro; IPR006953; Uso1_p115_head.
DR   Pfam; PF04871; Uso1_p115_C; 1.
DR   Pfam; PF04869; Uso1_p115_head; 1.
SQ   SEQUENCE   1171 AA;  131632 MW;  33DF50E5931ED060 CRC64;

  Query Match             55.7%;  Score 58.5;  DB 2;  Length 1171;
  Best Local Similarity   56.7%;  Pred. No. 1.2e+02;
  Matches   17;  Conservative    3;  Mismatches    5;  Indels    5;  Gaps    1;

Qy          1 AXAEA-----AEKAAKYAAEAAEKAAKAXA 25
              |:|||     |||||   |||| ||| |::
Db       1025 AEAEADAATGAEKAAAIEAEAAAKAAAAAS 1054


RESULT 34
Q9DF23
ID   Q9DF23      PRELIMINARY;      PRT;    92 AA.
AC   Q9DF23;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Skin-type antifreeze polypeptide AFP-2.
OS   Myoxocephalus scorpius (Shorthorn sculpin) (Daddy sculpin).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Scorpaeniformes;
OC   Cottoidei; Cottidae; Myoxocephalus.
OX   NCBI_TaxID=8097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RX   MEDLINE=98389738; PubMed=9722537; DOI=10.1074/jbc.273.36.23098;
RA   Low W.-K., Miao M., Ewart K.V., Yang D.S.C., Fletcher G.L., Hew C.L.;
RT   "Skin-type antifreeze protein from the shorthorn sculpin,
RT   Myoxocephalus scorpius. Expression and characterization of a Mr 9, 700
RT   recombinant protein.";
RL   J. Biol. Chem. 273:23098-23103(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RA   Low W.-K., Miao M., Ewart K.V., Yang D.S.C., Fletcher G.L., Hew C.L.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF305502; AAG25982.1; -.
DR   GO; GO:0005622; C:intracellular; IEA.
DR   GO; GO:0005840; C:ribosome; IEA.
DR   GO; GO:0050825; F:ice binding; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   GO; GO:0042309; P:homoiothermy; IEA.
DR   GO; GO:0006412; P:protein biosynthesis; IEA.
DR   GO; GO:0050826; P:response to freezing; IEA.
DR   InterPro; IPR000104; Antifreeze_1.
DR   InterPro; IPR001778; POA_allergen_C.
DR   InterPro; IPR001859; Ribosomal_P2.
DR   PRINTS; PR00308; ANTIFREEZEI.
DR   PRINTS; PR00833; POAALLERGEN.
DR   PRINTS; PR00456; RIBOSOMALP2.
SQ   SEQUENCE   92 AA;  7693 MW;  A3FCFD57B5CA8465 CRC64;

  Query Match             55.2%;  Score 58;  DB 2;  Length 92;
  Best Local Similarity   60.0%;  Pred. No. 17;
  Matches   15;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|:||| ||  || ||| ||  :|
Db          3 AAAKAAEAAAMAAANAAEAAATKAA 27


RESULT 35
Q7V6K8
ID   Q7V6K8      PRELIMINARY;      PRT;   124 AA.
AC   Q7V6K8;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Type I antifreeze protein.
GN   OrderedLocusNames=PMT1149;
OS   Prochlorococcus marinus (strain MIT 9313).
```

```
OC   Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;
OC   Prochlorococcus.
OX   NCBI_TaxID=74547;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22825698; PubMed=12917642; DOI=10.1038/nature01947;
RA   Rocap G., Larimer F.W., Lamerdin J.E., Malfatti S., Chain P.,
RA   Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA   Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,
RA   Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA   Webb E.A., Zinser E.R., Chisholm S.W.;
RT   "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT   niche differentiation.";
RL   Nature 424:1042-1047(2003).
DR   EMBL; BX572098; CAE21324.1; -.
KW   Complete proteome.
SQ   SEQUENCE   124 AA;  12053 MW;  0023FB3DBF04E16E CRC64;

  Query Match             55.2%;  Score 58;  DB 2;  Length 124;
  Best Local Similarity   57.6%;  Pred. No. 22;
  Matches   19;  Conservative    4;  Mismatches    2;  Indels    8;  Gaps    2;

Qy          1 AXAEAAEK------AAKYAAE--AAEKAAKAXA 25
              |:|||| |      ||| |||  ||:|||:|:|
Db         55 AAAEAAAKQAAEQAAAKAAAEAAAAKKAAEAAA 87


RESULT 36
O96113
ID   O96113      PRELIMINARY;      PRT;   387 AA.
AC   O96113;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Rifin.
GN   Name=PFB0035c;
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99021743; PubMed=9804551; DOI=10.1126/science.282.5391.1126;
RA   Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA   Koonin E.V., Shallom S., Mason T., Yu K., Fujii C., Pederson J.,
RA   Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
RA   Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA   Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT   "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT   falciparum.";
RL   Science 282:1126-1132(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA   Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA   Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA   Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA   Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
RA   Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA   Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA   McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA   Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA   Fraser C.M., Barrell B.;
RT   "Genome sequence of the human malaria parasite Plasmodium
RT   falciparum.";
RL   Nature 419:498-511(2002).
DR   EMBL; AE001367; AAC71797.2; -.
DR   PIR; A71625; A71625.
DR   InterPro; IPR011038; Calycin.
DR   InterPro; IPR006373; Rifin.
DR   InterPro; IPR002858; Rifin_STEVOR.
DR   Pfam; PF02009; Rifin_STEVOR; 1.
DR   TIGRFAMs; TIGR01477; RIFIN; 1.
SQ   SEQUENCE   387 AA;  42873 MW;  98846A6E588C2A35 CRC64;

  Query Match             55.2%;  Score 58;  DB 2;  Length 387;
  Best Local Similarity   56.5%;  Pred. No. 56;
  Matches   13;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              : | ||:||| ||:|||| |:
Db        304 IVEGAEQAAKAAAKAAEKGVTAA 326


RESULT 37
Q7RWB7
ID   Q7RWB7      PRELIMINARY;      PRT;   558 AA.
AC   Q7RWB7;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein (Probable proline-tRNA ligase).
GN   Name=NCU04449.1; Synonyms=G21B4.130;
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OR74A;
RA   Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA   Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA   Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA   Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA   Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA   Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA   Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA   Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA   Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA   Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
RA   DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA   Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA   Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA   Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT   "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL   Nature 0:0-0(2003).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA   Nyakatura G., Mewes H.W., Mannhaupt G.;
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   German Neurospora genome project;
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AABX01000782; EAA26671.1; -.
DR   EMBL; BX908808; CAF05998.1; -.
DR   HSSP; Q93N97; 1H4S.
DR   GO; GO:0005737; C:cytoplasm; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0004827; F:proline-tRNA ligase activity; IEA.
DR   GO; GO:0006433; P:prolyl-tRNA aminoacylation; IEA.
DR   GO; GO:0006412; P:protein biosynthesis; IEA.
DR   InterPro; IPR004154; HGTP_anticodon.
DR   InterPro; IPR004499; ProS_fam_I.
DR   InterPro; IPR002314; tRNA-synt_2b.
DR   InterPro; IPR002316; tRNA-synt_pro.
DR   InterPro; IPR006195; tRNA_ligase_II.
DR   Pfam; PF03129; HGTP_anticodon; 1.
DR   Pfam; PF00587; tRNA-synt_2b; 1.
DR   PRINTS; PR01046; TRNASYNTHPRO.
DR   TIGRFAMs; TIGR00408; proS_fam_I; 1.
```

```
DR   PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
KW   Hypothetical protein; Ligase.
SQ   SEQUENCE   558 AA;  63124 MW;  95E5E653E44E41A8 CRC64;

  Query Match             55.2%;  Score 58;  DB 2;  Length 558;
  Best Local Similarity   59.1%;  Pred. No. 75;
  Matches   13;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAK 22
              :: : |||||: ||: ||||||
Db         10 SALKKAEKAAQMAAKKAEKAAK 31


RESULT 38
Q6PF71
ID   Q6PF71      PRELIMINARY;      PRT;   575 AA.
AC   Q6PF71;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC057706; AAH57706.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR009007; Pept_Aspartic.
DR   InterPro; IPR003309; Treg_SCAN.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF02023; SCAN; 1.
DR   Pfam; PF00098; zf-CCHC; 1.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   PROSITE; PS50804; SCAN_BOX; 1.
DR   PROSITE; PS50158; ZF_CCHC; 1.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   575 AA;  62786 MW;  D60DF4183B8C81B6 CRC64;

  Query Match             55.2%;  Score 58;  DB 2;  Length 575;
  Best Local Similarity   57.7%;  Pred. No. 77;
  Matches   15;  Conservative    6;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          2 XAEAAEKAAKYAA--EAAEKAAKAXA 25
              :|||||:||:  |  ||||:||: :|
Db         49 QAEAAERAAERTAEREAAERAAERTA 74


RESULT 39
Q88YM9
ID   Q88YM9      PRELIMINARY;      PRT;   660 AA.
AC   Q88YM9;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   ABC transporter, ATP-binding protein.
GN   OrderedLocusNames=lp_0723;
OS   Lactobacillus plantarum.
OC   Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=1590;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCIMB 8826 / WCFS1;
RX   MEDLINE=22480296; PubMed=12566566; DOI=10.1073/pnas.0337704100;
RA   Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA   Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA   Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA   Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA   De Vos W.M., Siezen R.J.;
RT   "Complete genome sequence of Lactobacillus plantarum WCFS1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
CC   -!- SIMILARITY: Belongs to the ABC transporter family.
DR   EMBL; AL935253; CAD63322.1; -.
DR   HSSP; P58301; 1US8.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR   GO; GO:0000166; F:nucleotide binding; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003439; ABC_transporter.
DR   Pfam; PF00005; ABC_tran; 2.
DR   ProDom; PD000006; ABC_transporter; 2.
DR   SMART; SM00382; AAA; 2.
DR   PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
KW   ATP-binding; Complete proteome.
SQ   SEQUENCE   660 AA;  74176 MW;  10783CC765A085D1 CRC64;

  Query Match             55.2%;  Score 58;  DB 2;  Length 660;
  Best Local Similarity   63.6%;  Pred. No. 86;
  Matches   14;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 EAAEKAAKYAAEAAEKAAKAXA 25
              | || ||  |::||||||:|:|
Db        538 EQAEIAAAAASQAAEKAAEAGA 559


RESULT 40
P90534
ID   P90534      PRELIMINARY;      PRT;   809 AA.
AC   P90534;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
```

```
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Rsc12 (Fragment).
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX4;
RA   Iranfar N., Loomis W.F.;
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U83087; AAB40930.1; -.
DR   DictyBase; DDB0214855; rsc12.
DR   GO; GO:0005737; C:cytoplasm; IEA.
DR   GO; GO:0006817; P:phosphate transport; IEA.
DR   InterPro; IPR008160; Collagen.
DR   Pfam; PF01391; Collagen; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   809 AA;  80296 MW;  54B8FFF7BA344300 CRC64;

  Query Match             55.2%;  Score 58;  DB 2;  Length 809;
  Best Local Similarity   56.0%;  Pred. No. 1e+02;
  Matches   14;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              ::| ||::|| || ||:||| |:|
Db        765 STAAAAQQAAGAAAAAAQKAAMAAA 789


RESULT 41
POL1_TORVR
ID   POL1_TORVR     STANDARD;      PRT;  2197 AA.
AC   P29150; Q88875; Q88876;
DT   01-DEC-1992 (Rel. 24, Created)
DT   25-OCT-2004 (Rel. 45, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   RNA1 polyprotein (P1) [Contains: X1 protein; X2 protein; NTP-binding
DE   protein (NTB) (Membrane-binding protein); Viral genome-linked protein
DE   (VPg); 3C-like protease (EC 3.4.22.-) (PRO); RNA-directed RNA
DE   polymerase (EC 2.7.7.48) (POL)].
OS   Tomato ringspot virus (subgroup C / isolate raspberry) (ToRSV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Comoviridae;
OC   Nepovirus.
OX   NCBI_TaxID=12281;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95146991; PubMed=7844569;
RA   Rott M.E., Gilchrist A., Lee L., Rochon D.M.;
RT   "Nucleotide sequence of tomato ringspot virus RNA1.";
RL   J. Gen. Virol. 76:465-473(1995).
RN   [2]
RP   SEQUENCE OF 1-354 FROM N.A.
RX   MEDLINE=92024112; PubMed=1926788;
RA   Rott M.E., Tremaine J.H., Rochon D.M.;
RT   "Comparison of the 5' and 3' termini of tomato ringspot virus RNA1 and
RT   RNA2: evidence for RNA recombination.";
RL   Virology 185:468-472(1991).
RN   [3]
RP   CHARACTERIZATION OF PROTEASE, AND MUTAGENESIS OF HIS-1283 AND
RP   HIS-1451.
RX   PubMed=9049338;
RA   Hans F., Sanfacon H.;
RT   "Tomato ringspot nepovirus protease: characterization and cleavage
RT   site specificity.";
RL   J. Gen. Virol. 76:917-927(1995).
RN   [4]
RP   SEQUENCE OF 1213-1247 AND 1487-1502, PROCESSING OF POLYPROTEIN, AND
RP   MUTAGENESIS OF GLN-1465 AND GLN-1486.
RX   PubMed=10092022;
RA   Wang A., Carrier K., Chisholm J., Wieczorek A., Huguenot C.,
RA   Sanfacon H.;
RT   "Proteolytic processing of tomato ringspot nepovirus 3C-like protease
RT   precursors: definition of the domains for the VPg, protease and
RT   putative RNA-dependent RNA polymerase.";
RL   J. Gen. Virol. 80:799-809(1999).
RN   [5]
RP   PROCESSING OF POLYPROTEIN, AND MUTAGENESIS OF GLN-423; GLN-620;
RP   GLN-1212 AND HIS-1283.
RX   PubMed=11038391;
RA   Wang A., Sanfacon H.;
RT   "Proteolytic processing at a novel cleavage site in the N-terminal
RT   region of the tomato ringspot nepovirus RNA-1-encoded polyprotein in
RT   vitro.";
RL   J. Gen. Virol. 81:2771-2781(2000).
RN   [6]
RP   SUBCELLULAR LOCATION OF THE NBT-VPG PROTEIN.
RX   PubMed=12477857; DOI=10.1128/JVI.77.1.523-534.2003;
RA   Han S., Sanfacon H.;
RT   "Tomato ringspot virus proteins containing the nucleoside triphosphate
RT   binding domain are transmembrane proteins that associate with the
RT   endoplasmic reticulum and cofractionate with replication complexes.";
RL   J. Virol. 77:523-534(2003).
RN   [7]
RP   TOPOLOGY OF THE NBT-VPG PROTEIN, GLYCOSYLATION, AND MUTAGENESIS OF
RP   THR-1230.
RX   PubMed=14769910; DOI=10.1099/vir.0.19612-0;
RA   Wang A., Han S., Sanfacon H.;
RT   "Topogenesis in membranes of the NTB-VPg protein of Tomato ringspot
RT   nepovirus: definition of the C-terminal transmembrane domain.";
RL   J. Gen. Virol. 85:535-545(2004).
CC   -!- FUNCTION: The 3C-like protease is a thiol protease that cleaves at
CC       Gln-|-Gly or Gln-|-Ser sites in the P1 and P2 polyproteins.
CC   -!- FUNCTION: The VPg-NBT polyprotein may act as a membrane-anchor for
CC       the replication complex.
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- SUBCELLULAR LOCATION: The NTB-VPg polyprotein is associated with
CC       endoplasmic-derived membranes that are active in viral
CC       replication. VPg localizes to the endoplasmic reticulum lumen. NTB
CC       is an integral membrane protein.
CC   -!- PTM: Specific enzymatic cleavages by 3C-like protease in vivo
CC       yield mature proteins. 3C-like protease is autocatalytically
CC       processed. NBT exists as NBT-VPg polyprotein as well as NBT mature
CC       protein.
CC   -!- PTM: VPg is covalently linked to the 5' end of genomic RNA (By
CC       similarity).
CC   -!- SIMILARITY: Belongs to the nepoviruses RNA1 polyprotein family.
CC   -!- SIMILARITY: Contains 1 peptidase C3 domain.
CC   -!- CAUTION: It is uncertain whether Met-1 or Met-122 is the
CC       initiator.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L19655; AAA78254.1; -.
DR   EMBL; M73822; AAA47941.1; -.
DR   EMBL; M73822; AAA47942.1; ALT_INIT.
DR   PIR; A40787; GNVVSR.
DR   MEROPS; C03.012; -.
DR   InterPro; IPR004004; Pept_Calici.
DR   InterPro; IPR010464; Picornain_3C.
DR   InterPro; IPR000605; RNA_helicase.
DR   InterPro; IPR007095; RNA_pol_DS_PS.
DR   InterPro; IPR001205; RNA_pol_P3D.
DR   InterPro; IPR007094; RNA_pol_PSvir.
DR   Pfam; PF06343; Picornain_3C; 1.
DR   Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR   Pfam; PF00910; RNA_helicase; 1.
DR   PRINTS; PR00918; CALICVIRUSNS.
KW   ATP-binding; Covalent protein-RNA linkage; Direct protein sequencing;
KW   Glycoprotein; Hydrolase; Polyprotein; Protease;
KW   RNA-directed RNA polymerase; Thiol protease; Transferase;
```

```
KW   Transmembrane.
FT   CHAIN         1    423       X1 protein.
FT   CHAIN       424    620       X2 protein.
FT   CHAIN       621   1212       NTP-binding protein.
FT   CHAIN      1213   1239       Viral genome-linked protein.
FT   CHAIN      1240   1486       3C-like protease.
FT   CHAIN      1487   2197       RNA-directed RNA polymerase.
FT   DOMAIN      621   1167       Cytoplasmic.
FT   TRANSMEM   1168   1188       Probable.
FT   DOMAIN     1189   1212       Lumenal.
FT   NP_BIND     796    803       ATP (Potential).
FT   DOMAIN      149    152       Poly-Pro.
FT   DOMAIN      230    235       Poly-Pro.
FT   ACT_SITE   1283   1283       3C-like protease (Probable).
FT   ACT_SITE   1331   1331       3C-like protease (Potential).
FT   ACT_SITE   1433   1433       3C-like protease (Potential).
FT   SITE       1451   1451       Involved in the cleavage site
FT                                specificity.
FT   CARBOHYD   1228   1228       N-linked (GlcNAc. . .).
FT   MUTAGEN     423    423       Missing: No cleavage between X1 and X2.
FT   MUTAGEN     620    620       Missing: No cleavage between X2 and NTB.
FT   MUTAGEN    1212   1212       Missing: No cleavage NTB and VPg.
FT   MUTAGEN    1230   1230       T->A: Complete loss of N-linked
FT                                glycosylation.
FT   MUTAGEN    1283   1283       H->D: Complete loss of protease activity.
FT   MUTAGEN    1451   1451       H->L: Complete loss of protease activity.
FT   MUTAGEN    1465   1465       Q->A: No effect.
FT   MUTAGEN    1486   1486       Q->A: No cleavage between 3C-like
FT                                protease and RNA-directed RNA polymerase.
FT   CONFLICT   1230   1230       T -> A (in Ref. 4; AA sequence).
SQ   SEQUENCE   2197 AA;  244128 MW;  2D8EF928E58DBC89 CRC64;

  Query Match             55.2%;  Score 58;  DB 1;  Length 2197;
  Best Local Similarity   70.0%;  Pred. No. 2.3e+02;
  Matches   14;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          6 AEKAAKYAAEAAEKAAKAXA 25
              | ||||||| || | | |:|
Db        180 ARKAAKYAAFAARKKAAAVA 199


RESULT 42
Q89IE3
ID   Q89IE3      PRELIMINARY;      PRT;   638 AA.
AC   Q89IE3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Bll5696 protein.
GN   OrderedLocusNames=bll5696;
OS   Bradyrhizobium japonicum.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Bradyrhizobium.
OX   NCBI_TaxID=375;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=USDA110;
RX   MEDLINE=22484998; PubMed=12597275;
RA   Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA   Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA   Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA   Tabata S.;
RT   "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT   Bradyrhizobium japonicum USDA110.";
RL   DNA Res. 9:189-197(2002).
DR   EMBL; AP005955; BAC50961.1; -.
DR   GO; GO:0030693; F:caspase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001309; ICE_p20.
DR   PROSITE; PS50208; CASPASE_P20; 1.
KW   Complete proteome.
SQ   SEQUENCE   638 AA;  68387 MW;  9519C8A749528B5B CRC64;

  Query Match             54.8%;  Score 57.5;  DB 2;  Length 638;
  Best Local Similarity   53.3%;  Pred. No. 95;
  Matches   16;  Conservative    4;  Mismatches    5;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEK-----AAKYAAEAAEKAAKAXA 25
              |: :||||     ||| ||| ||: || :|
Db        415 AEKQAAEKAKAELAAKQAAEKAEQPAKPTA 444


RESULT 43
O15860
ID   O15860      PRELIMINARY;      PRT;   190 AA.
AC   O15860;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Merozoite surface protein 3 (Fragment).
GN   Name=SPAM;
OS   Plasmodium falciparum.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5833;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97391128; PubMed=9247935; DOI=10.1016/S0166-6851(97)00067-4;
RA   Huber W., Felger I., Matile H., Lipps H.J., Steiger S., Beck H.P.;
RT   "Limited sequence polymorphism in the Plasmodium falciparum merozoite
RT   surface protein 3.";
RL   Mol. Biochem. Parasitol. 87:231-234(1997).
DR   EMBL; AF001149; AAC47674.1; -.
DR   InterPro; IPR010784; Merozoite_SPAM.
DR   Pfam; PF07133; Merozoite_SPAM; 1.
KW   Merozoite.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21170 MW;  9BD627A8758AC41C CRC64;

  Query Match             54.3%;  Score 57;  DB 2;  Length 190;
  Best Local Similarity   56.5%;  Pred. No. 41;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|| |||||: | :|||:| :|:
Db         71 AAEEAEKAAEEAEKAAEQAEQAS 93


RESULT 44
Q8X965
ID   Q8X965      PRELIMINARY;      PRT;   394 AA.
AC   Q8X965; Q7AGI8;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Membrane spanning protein, required for outer membrane integrity
DE   (Membrane spanning protein TolA).
GN   Name=tolA; OrderedLocusNames=ECs0774, z0907;
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927 /  EHEC;
RX   MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=O157:H7 / RIMD 0509952 / EHEC;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
DR   EMBL; AE005252; AAG55075.1; -.
DR   EMBL; AP002553; BAB34197.1; -.
DR   PIR; F90725; F90725.
DR   PIR; G85576; G85576.
DR   HSSP; P19934; 1TOL.
DR   InterPro; IPR010528; TolA.
DR   Pfam; PF06519; TolA; 1.
KW   Complete proteome.
SQ   SEQUENCE   394 AA;  40517 MW;  5B58D8E8230BDE28 CRC64;

  Query Match             54.3%;  Score 57;  DB 2;  Length 394;
  Best Local Similarity   56.0%;  Pred. No. 74;
  Matches   14;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |: :|||:||| ||  |:| |:|:|
Db        151 ADDKAAEEAAKKAAADAKKKAEAEA 175


RESULT 45
Q875A8
ID   Q875A8      PRELIMINARY;      PRT;   508 AA.
AC   Q875A8;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Similar to Danio rerio protein-kinase.
OS   Podospora anserina.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetidae; Sordariales; Lasiosphaeriaceae; Podospora.
OX   NCBI_TaxID=5145;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Genoscope;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX088700; CAD60707.1; -.
DR   HSSP; P08622; 1BQZ.
DR   GO; GO:0016301; F:kinase activity; IEA.
DR   InterPro; IPR001623; DnaJ_N.
DR   InterPro; IPR008940; Prenyl_trans.
DR   InterPro; IPR001440; TPR.
DR   Pfam; PF00226; DnaJ; 1.
DR   Pfam; PF00515; TPR_1; 1.
DR   SMART; SM00271; DnaJ; 1.
DR   SMART; SM00028; TPR; 3.
DR   PROSITE; PS50076; DNAJ_2; 1.
DR   PROSITE; PS50005; TPR; 2.
DR   PROSITE; PS50293; TPR_REGION; 1.
KW   Kinase; Repeat; TPR repeat.
SQ   SEQUENCE   508 AA;  55747 MW;  5826AC06E1E8DB9C CRC64;

  Query Match             54.3%;  Score 57;  DB 2;  Length 508;
  Best Local Similarity   70.0%;  Pred. No. 91;
  Matches   14;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 AAEKAAKYAAEAAEKAAKAX 24
              || :||| ||:||| ||||:
Db        136 AAVEAAKAAAKAAEAAAKAG 155


Search completed: July 11, 2005, 09:46:25
Job time : 170 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2005, 09:22:59 ; Search time 161 Seconds
(without alignments)
60.056 Million cell updates/sec

Title: SEQ1
Perfect score: 105
Sequence: 1 axaeaaekaakyaaeaaekaakaxa 25

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	105	100.0	25	4	AAB66787	Aab66787 Amphipath
2	67	63.8	104	7	ADE10685	Ade10685 Structura
3	67	63.8	104	7	ADE10635	Ade10635 Structura
4	67	63.8	104	8	ADK15654	Adk15654 Nucleatin
5	67	63.8	104	8	ADK15704	Adk15704 Library f
6	65.5	62.4	428	6	ABU27824	Abu27824 Protein e
7	64	61.0	28	8	ADO43180	Ado43180 Peptide u
8	64	61.0	28	8	ADO43177	Ado43177 Peptide u
9	64	61.0	104	7	ADE10683	Ade10683 Structura
10	64	61.0	104	7	ADE10682	Ade10682 Structura
11	64	61.0	104	7	ADE10633	Ade10633 Structura
12	64	61.0	104	7	ADE10632	Ade10632 Structura
13	64	61.0	104	8	ADK15652	Adk15652 Nucleatin
14	64	61.0	104	8	ADK15701	Adk15701 Library f
15	64	61.0	104	8	ADK15702	Adk15702 Library f
16	64	61.0	104	8	ADK15651	Adk15651 Nucleatin
17	63.5	60.5	104	7	ADE10684	Ade10684 Structura
18	63.5	60.5	104	7	ADE10634	Ade10634 Structura
19	63.5	60.5	104	8	ADK15703	Adk15703 Library f
20	63.5	60.5	104	8	ADK15653	Adk15653 Nucleatin
21	63	60.0	59	7	ADE10698	Ade10698 Structura
22	63	60.0	59	7	ADE10648	Ade10648 Structura
23	63	60.0	59	8	ADK15717	Adk15717 Library f
24	63	60.0	59	8	ADK15667	Adk15667 Nucleatin
25	63	60.0	67	7	ADE10697	Ade10697 Structura
26	63	60.0	67	7	ADE10647	Ade10647 Structura
27	63	60.0	67	8	ADK15666	Adk15666 Nucleatin
28	63	60.0	67	8	ADK15716	Adk15716 Library f
29	63	60.0	75	7	ADE10696	Ade10696 Structura
30	63	60.0	75	7	ADE10646	Ade10646 Structura
31	63	60.0	75	8	ADK15715	Adk15715 Library f
32	63	60.0	75	8	ADK15665	Adk15665 Nucleatin
33	63	60.0	83	7	ADE10695	Ade10695 Structura
34	63	60.0	83	7	ADE10645	Ade10645 Structura
35	63	60.0	83	8	ADK15714	Adk15714 Library f
36	63	60.0	83	8	ADK15664	Adk15664 Nucleatin
37	63	60.0	88	7	ADE10642	Ade10642 Structura
38	63	60.0	88	7	ADE10692	Ade10692 Structura
39	63	60.0	88	8	ADK15711	Adk15711 Library f
40	63	60.0	88	8	ADK15661	Adk15661 Nucleatin
41	63	60.0	91	7	ADE10694	Ade10694 Structura
42	63	60.0	91	7	ADE10644	Ade10644 Structura
43	63	60.0	91	8	ADK15663	Adk15663 Nucleatin
44	63	60.0	91	8	ADK15713	Adk15713 Library f
45	63	60.0	104	7	ADE10690	Ade10690 Structura

ALIGNMENTS

RESULT 1
AAB66787
ID AAB66787 standard; peptide; 25 AA.
XX
AC AAB66787;
XX
DT 11-APR-2001 (first entry)
XX
DE Amphipathic peptide conjugate.
XX
KW Amphipathic; lipid bilayer; detergent.
XX
OS Synthetic.
XX
PN WO200102425-A2.
XX
PD 11-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-CA000773.
XX
PR 29-JUN-1999; 99US-0140988P.
XX
PA (UYHE-) UNIV HEALTH NETWORK.
XX
PI Prive G;
XX
DR WPI; 2001-138120/14.
XX
PT New amphiphatic peptide conjugate having detergent properties, and
PT hydrophobic and hydrophilic phase, useful e.g. for stabilizing and
PT crystallizing proteins and membrane proteins, as cytolytic agents,
PT surfactants or emulsifiers.
XX
PS Claim 1; Page 22; 29pp; English.
XX
CC The present invention relates to an amphipathic peptide conjugate having
CC detergent properties and a hydrophobic and hydrophilic face. The
CC amphipathic peptide conjugate may be used for the stabilization and
CC crystallization of proteins and membrane proteins,for modifying the
CC properties of lipid bilayer membranes, as cytolytic agents, as molecules
CC that can facilitate the transport of polar molecules across biological
CC membranes, and as emulsifiers and surfactants
XX
SQ Sequence 25 AA;

Query Match 100.0%; Score 105; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.2e-07;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |||||||||||||||||||||||||
Db          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
```

RESULT 2
ADE10685
ID ADE10685 standard; protein; 104 AA.
XX
AC ADE10685;
XX
DT 29-JAN-2004 (first entry)
XX
DE Structurally biased random peptide library scaffold protein seqid 92.
XX
KW fusion nucleic acid library; scaffold protein; bioactive peptide;
KW phenotype change; cell morphology; cell growth; cell viability;
KW cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW loss of cell division; decreased cell growth; brca-1; brca-2;
KW tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW skin biology; cosmeceutical; endocrinology; infectious disease;
KW drug toxicity; drug resistance; inflammation; allergic response;
KW scaffold protein.
XX
OS Synthetic.
XX
PN US2003143562-A1.
XX
PD 31-JUL-2003.
XX
PF 20-JUN-2002; 2002US-00177725.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2003-829786/77.
XX
PT Novel library of fusion nucleic acids each of which has fused first and
PT second nucleic acids encoding scaffold protein and library peptide having
PT alpha helical biasing sequence, respectively, useful in screening
PT methods.
XX
PS Disclosure; SEQ ID NO 92; 110pp; English.
XX
CC The invention describes a library (I) of fusion nucleic acids, where each
CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC library peptide sequence comprising an alpha helical biasing sequence;
CC where N1 is fused to N2. Disclosed is a method for screening bioactive
CC peptides conferring a change in specific phenotype such as cell
CC morphology, cell growth, cell viability, adhesion to substrates or other
CC cells, and cellular density; changes in the expression of one or more
CC RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC peptide identified by above mentioned method is used to generate more
CC candidate peptides and to identify target molecules, i.e., the molecules
CC with which the bioactive peptide interacts. The peptide(s) can be
CC combined with other pharmacologic activators to study the epistatic
CC relationships of signal transduction pathways in question. The disclosed
CC method is also useful in cancer applications. Random libraries can be
CC introduced into any tumour cell (primary or cultured), and peptides
CC identified which by themselves induce apoptosis, cell death, loss of cell
CC division or decreased cell growth. The method is also useful for
CC screening of bioactive peptides which restore the constitutive function
CC of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC important in breast cancer such as the adenomatous polyposis coli gene
CC (APC) and the Drosophila discs-large gene (DIg), which are components of
CC cell-cell junctions. The methods are useful in cardiovascular
CC applications, neurobiology applications, bone biology applications, skin
CC biology applications, cosmeceutical applications, endocrinology
CC applications, infectious disease applications, drug toxicities and drug
CC resistance applications, immunobiology, inflammation, and allergic
CC response applications, and biotechnology applications. The peptide
CC library can easily be monitored, both for its presence within cells and
CC its quantity. The expression of structurally biased libraries generate
CC elevated cellular concentration of peptides having a given structural
CC bias and thus increase the hit rate for targets that bind such
CC structures. This is the amino acid sequence of a scaffold protein used in
CC peptide libraries ot hold the library peptide in a conformationally
CC restricted form.
XX
SQ Sequence 104 AA;

```
  Query Match             63.8%;  Score 67;  DB 7;  Length 104;
  Best Local Similarity   68.0%;  Pred. No. 0.18;
  Matches   17;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||| ||| ||||| |||  :|
Db         10 AAAAAAEAAAKAAAEAAAKAAAEAA 34
```

RESULT 3
ADE10635
ID ADE10635 standard; protein; 104 AA.
XX
AC ADE10635;
XX
DT 29-JAN-2004 (first entry)
XX
DE Structurally biased random peptide library related protein seqid 42.
XX
KW fusion nucleic acid library; scaffold protein; bioactive peptide;
KW phenotype change; cell morphology; cell growth; cell viability;
KW cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW loss of cell division; decreased cell growth; brca-1; brca-2;
KW tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW skin biology; cosmeceutical; endocrinology; infectious disease;
KW drug toxicity; drug resistance; inflammation; allergic response.
XX
OS Synthetic.
XX
PN US2003143562-A1.
XX
PD 31-JUL-2003.
XX
PF 20-JUN-2002; 2002US-00177725.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2003-829786/77.
XX
PT Novel library of fusion nucleic acids each of which has fused first and
PT second nucleic acids encoding scaffold protein and library peptide having
PT alpha helical biasing sequence, respectively, useful in screening
PT methods.
XX
PS Example 6; SEQ ID NO 42; 110pp; English.
XX
CC The invention describes a library (I) of fusion nucleic acids, where each
CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC library peptide sequence comprising an alpha helical biasing sequence;

```
CC   where N1 is fused to N2. Disclosed is a method for screening bioactive
CC   peptides conferring a change in specific phenotype such as cell
CC   morphology, cell growth, cell viability, adhesion to substrates or other
CC   cells, and cellular density; changes in the expression of one or more
CC   RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC   in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC   lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC   peptide identified by above mentioned method is used to generate more
CC   candidate peptides and to identify target molecules, i.e., the molecules
CC   with which the bioactive peptide interacts. The peptide(s) can be
CC   combined with other pharmacologic activators to study the epistatic
CC   relationships of signal transduction pathways in question. The disclosed
CC   method is also useful in cancer applications. Random libraries can be
CC   introduced into any tumour cell (primary or cultured), and peptides
CC   identified which by themselves induce apoptosis, cell death, loss of cell
CC   division or decreased cell growth. The method is also useful for
CC   screening of bioactive peptides which restore the constitutive function
CC   of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC   important in breast cancer such as the adenomatous polyposis coli gene
CC   (APC) and the Drosophila discs-large gene (DIg), which are components of
CC   cell-cell junctions. The methods are useful in cardiovascular
CC   applications, neurobiology applications, bone biology applications, skin
CC   biology applications, cosmeceutical applications, endocrinology
CC   applications, infectious disease applications, drug toxicities and drug
CC   resistance applications, immunobiology, inflammation, and allergic
CC   response applications, and biotechnology applications. The peptide
CC   library can easily be monitored, both for its presence within cells and
CC   its quantity. The expression of structurally biased libraries generate
CC   elevated cellular concentration of peptides having a given structural
CC   bias and thus increase the hit rate for targets that bind such
CC   structures. This is the amino acid sequence of a protein associated with
CC   fused nucleic acid and random peptide libraries of the invention.
XX
SQ   Sequence 104 AA;

  Query Match             63.8%;  Score 67;  DB 7;  Length 104;
  Best Local Similarity   68.0%;  Pred. No. 0.18;
  Matches   17;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||| ||| ||||| |||  :|
Db         10 AAAAAAEAAAKAAAEAAAKAAAEAA 34


RESULT 4
ADK15654
ID   ADK15654 standard; peptide; 104 AA.
XX
AC   ADK15654;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Nucleating sequence-containing library fusion protein #36.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2004-033956/03.
XX
PT   Library of fusion polypeptides in which each polypeptides comprises
PT   scaffold protein and library peptide having alpha helical biasing
PT   sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS   Example 6; SEQ ID NO 42; 110pp; English.
XX
CC   The invention comprises a library of fusion nucleic acids, where each
CC   encoded protein contains a scaffold protein (e.g. a green fluorescent
CC   protein - GFP) and a library peptide sequence comprising an alpha helical
CC   biasing sequence, or a scaffold protein, a library peptide and a
CC   nucleating sequence. The library of the invention is useful for screening
CC   bioactive peptides conferring a particular phenotype. The present amino
CC   acid sequence represents a library protein containing a nucleating
CC   sequence.
XX
SQ   Sequence 104 AA;

  Query Match             63.8%;  Score 67;  DB 8;  Length 104;
  Best Local Similarity   68.0%;  Pred. No. 0.18;
  Matches   17;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||| ||| ||||| |||  :|
Db         10 AAAAAAEAAAKAAAEAAAKAAAEAA 34


RESULT 5
ADK15704
ID   ADK15704 standard; peptide; 104 AA.
XX
AC   ADK15704;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Library fusion protein-related scaffold protein #36.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2004-033956/03.
XX
PT   Library of fusion polypeptides in which each polypeptides comprises
PT   scaffold protein and library peptide having alpha helical biasing
PT   sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS   Disclosure; SEQ ID NO 92; 110pp; English.
XX
CC   The invention comprises a library of fusion nucleic acids, where each
CC   encoded protein contains a scaffold protein (e.g. a green fluorescent
```

```
CC   protein - GFP) and a library peptide sequence comprising an alpha helical
CC   biasing sequence, or a scaffold protein, a library peptide and a
CC   nucleating sequence. The library of the invention is useful for screening
CC   bioactive peptides conferring a particular phenotype. The present amino
CC   acid sequence represents a scaffold protein.
XX
SQ   Sequence 104 AA;

  Query Match             63.8%;  Score 67;  DB 8;  Length 104;
  Best Local Similarity   68.0%;  Pred. No. 0.18;
  Matches   17;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||| ||| ||||| |||  :|
Db         10 AAAAAAEAAAKAAAEAAAKAAAEAA 34


RESULT 6
ABU27824
ID   ABU27824 standard; protein; 428 AA.
XX
AC   ABU27824;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #13351.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Enterobacter cloacae.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,  Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,  Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA31694.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 25; SEQ ID NO 55748; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC   the target prokaryotic essential genes. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 428 AA;

  Query Match             62.4%;  Score 65.5;  DB 6;  Length 428;
  Best Local Similarity   62.1%;  Pred. No. 1.4;
  Matches   18;  Conservative    4;  Mismatches    2;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAA-----KYAAEAAEKAAKAX 24
              |:||||:|||     | |||||:||| |:
Db        210 AEAEAAKKAAQEAEKKAAAEAAKKAAAAE 238


RESULT 7
ADO43180
ID   ADO43180 standard; peptide; 28 AA.
XX
AC   ADO43180;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Peptide used for coded probe synthesis.
XX
KW   Nano-barcode; scanning probe microscopy; probe.
XX
OS   Synthetic.
XX
PN   WO2004038037-A2.
XX
PD   06-MAY-2004.
XX
PF   22-SEP-2003; 2003WO-US029726.
XX
PR   20-SEP-2002; 2002US-00251152.
PR   19-SEP-2003; 2003US-00667004.
XX
PA   (ITLC ) INTEL CORP.
XX
PI   Chan S,  Su X,  Yamakawa M;
XX
DR   WPI; 2004-399960/37.
XX
PT   Detecting, identifying and sequencing of biomolecules using controlled
PT   alignment of nano-barcodes encoding specific information for scanning
PT   probe microscopy, useful in the fields of molecular biology.
XX
PS   Example 2; Page 44; 63pp; English.
XX
CC   The present sequence is that of a peptide of potential use for production
CC   of a coded probe useful in the method of the invention. The invention
CC   provides methods, apparatus and compositions for the detection,
CC   identification and/or sequencing of biomolecules, such as nucleic acids
CC   or proteins. Coded probes comprising a probe molecule attached to one or
CC   more nano-barcodes are allowed to bind to target molecule(s). After
CC   binding and separation from unbound coded probes, the bound coded probes
CC   are aligned on a surface and analysed by scanning probe microscopy (SPM).
CC   The methods allow the sequencing of long nucleic acid sequences in a
CC   single sequencing run, high speed of obtaining sequence data, low cost of
CC   sequencing and high efficiency in terms of operator time, and sensitive
```

```
CC   and accurate detection and/or identification of nucleic acids with low
CC   incidence of false positive results.
XX
SQ   Sequence 28 AA;

  Query Match             61.0%;  Score 64;  DB 8;  Length 28;
  Best Local Similarity   65.2%;  Pred. No. 0.1;
  Matches   15;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|||| :||: |||||| ||:|:
Db          1 AAEAAAEAAEAAAEAAEAAAEAA 23


RESULT 8
ADO43177
ID   ADO43177 standard; peptide; 28 AA.
XX
AC   ADO43177;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Peptide used for coded probe synthesis.
XX
KW   Nano-barcode; scanning probe microscopy; probe.
XX
OS   Synthetic.
XX
PN   WO2004038037-A2.
XX
PD   06-MAY-2004.
XX
PF   22-SEP-2003; 2003WO-US029726.
XX
PR   20-SEP-2002; 2002US-00251152.
PR   19-SEP-2003; 2003US-00667004.
XX
PA   (ITLC ) INTEL CORP.
XX
PI   Chan S,  Su X,  Yamakawa M;
XX
DR   WPI; 2004-399960/37.
XX
PT   Detecting, identifying and sequencing of biomolecules using controlled
PT   alignment of nano-barcodes encoding specific information for scanning
PT   probe microscopy, useful in the fields of molecular biology.
XX
PS   Example 2; Page 44; 63pp; English.
XX
CC   The present sequence is that of a peptide of potential use for production
CC   of a coded probe useful in the method of the invention. The invention
CC   provides methods, apparatus and compositions for the detection,
CC   identification and/or sequencing of biomolecules, such as nucleic acids
CC   or proteins. Coded probes comprising a probe molecule attached to one or
CC   more nano-barcodes are allowed to bind to target molecule(s). After
CC   binding and separation from unbound coded probes, the bound coded probes
CC   are aligned on a surface and analysed by scanning probe microscopy (SPM).
CC   The methods allow the sequencing of long nucleic acid sequences in a
CC   single sequencing run, high speed of obtaining sequence data, low cost of
CC   sequencing and high efficiency in terms of operator time, and sensitive
CC   and accurate detection and/or identification of nucleic acids with low
CC   incidence of false positive results.
XX
SQ   Sequence 28 AA;

  Query Match             61.0%;  Score 64;  DB 8;  Length 28;
  Best Local Similarity   65.2%;  Pred. No. 0.1;
  Matches   15;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|:|| |||| ||:||: ||||:
Db          1 AAKAAAKAAKAAAKAAKAAAKAA 23
```

```
RESULT 9
ADE10683
ID   ADE10683 standard; protein; 104 AA.
XX
AC   ADE10683;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Structurally biased random peptide library scaffold protein seqid 90.
XX
KW   fusion nucleic acid library; scaffold protein; bioactive peptide;
KW   phenotype change; cell morphology; cell growth; cell viability;
KW   cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW   loss of cell division; decreased cell growth; brca-1; brca-2;
KW   tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW   Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW   skin biology; cosmeceutical; endocrinology; infectious disease;
KW   drug toxicity; drug resistance; inflammation; allergic response;
KW   scaffold protein.
XX
OS   Synthetic.
XX
PN   US2003143562-A1.
XX
PD   31-JUL-2003.
XX
PF   20-JUN-2002; 2002US-00177725.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2003-829786/77.
XX
PT   Novel library of fusion nucleic acids each of which has fused first and
PT   second nucleic acids encoding scaffold protein and library peptide having
PT   alpha helical biasing sequence, respectively, useful in screening
PT   methods.
XX
PS   Disclosure; SEQ ID NO 90; 110pp; English.
XX
CC   The invention describes a library (I) of fusion nucleic acids, where each
CC   fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC   scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC   library peptide sequence comprising an alpha helical biasing sequence;
CC   where N1 is fused to N2. Disclosed is a method for screening bioactive
CC   peptides conferring a change in specific phenotype such as cell
CC   morphology, cell growth, cell viability, adhesion to substrates or other
CC   cells, and cellular density; changes in the expression of one or more
CC   RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC   in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC   lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC   peptide identified by above mentioned method is used to generate more
CC   candidate peptides and to identify target molecules, i.e., the molecules
CC   with which the bioactive peptide interacts. The peptide(s) can be
CC   combined with other pharmacologic activators to study the epistatic
CC   relationships of signal transduction pathways in question. The disclosed
CC   method is also useful in cancer applications. Random libraries can be
CC   introduced into any tumour cell (primary or cultured), and peptides
CC   identified which by themselves induce apoptosis, cell death, loss of cell
CC   division or decreased cell growth. The method is also useful for
CC   screening of bioactive peptides which restore the constitutive function
CC   of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC   important in breast cancer such as the adenomatous polyposis coli gene
CC   (APC) and the Drosophila discs-large gene (DIg), which are components of
CC   cell-cell junctions. The methods are useful in cardiovascular
CC   applications, neurobiology applications, bone biology applications, skin
CC   biology applications, cosmeceutical applications, endocrinology
CC   applications, infectious disease applications, drug toxicities and drug
```

```
CC   resistance applications, immunobiology, inflammation, and allergic
CC   response applications, and biotechnology applications. The peptide
CC   library can easily be monitored, both for its presence within cells and
CC   its quantity. The expression of structurally biased libraries generate
CC   elevated cellular concentration of peptides having a given structural
CC   bias and thus increase the hit rate for targets that bind such
CC   structures. This is the amino acid sequence of a scaffold protein used in
CC   peptide libraries ot hold the library peptide in a conformationally
CC   restricted form.
XX
SQ   Sequence 104 AA;

  Query Match             61.0%;  Score 64;  DB 7;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.45;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|||| |||  || ||| ||||:|
Db          9 AAAEAAAKAA--AAAAAEAAAKAAA 31


RESULT 10
ADE10682
ID   ADE10682 standard; protein; 104 AA.
XX
AC   ADE10682;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Structurally biased random peptide library scaffold protein seqid 89.
XX
KW   fusion nucleic acid library; scaffold protein; bioactive peptide;
KW   phenotype change; cell morphology; cell growth; cell viability;
KW   cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW   loss of cell division; decreased cell growth; brca-1; brca-2;
KW   tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW   Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW   skin biology; cosmeceutical; endocrinology; infectious disease;
KW   drug toxicity; drug resistance; inflammation; allergic response;
KW   scaffold protein.
XX
OS   Synthetic.
XX
PN   US2003143562-A1.
XX
PD   31-JUL-2003.
XX
PF   20-JUN-2002; 2002US-00177725.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2003-829786/77.
XX
PT   Novel library of fusion nucleic acids each of which has fused first and
PT   second nucleic acids encoding scaffold protein and library peptide having
PT   alpha helical biasing sequence, respectively, useful in screening
PT   methods.
XX
PS   Disclosure; SEQ ID NO 89; 110pp; English.
XX
CC   The invention describes a library (I) of fusion nucleic acids, where each
CC   fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC   scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC   library peptide sequence comprising an alpha helical biasing sequence;
CC   where N1 is fused to N2. Disclosed is a method for screening bioactive
CC   peptides conferring a change in specific phenotype such as cell
CC   morphology, cell growth, cell viability, adhesion to substrates or other
CC   cells, and cellular density; changes in the expression of one or more
CC   RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC   in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC   lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC   peptide identified by above mentioned method is used to generate more
CC   candidate peptides and to identify target molecules, i.e., the molecules
CC   with which the bioactive peptide interacts. The peptide(s) can be
CC   combined with other pharmacologic activators to study the epistatic
CC   relationships of signal transduction pathways in question. The disclosed
CC   method is also useful in cancer applications. Random libraries can be
CC   introduced into any tumour cell (primary or cultured), and peptides
CC   identified which by themselves induce apoptosis, cell death, loss of cell
CC   division or decreased cell growth. The method is also useful for
CC   screening of bioactive peptides which restore the constitutive function
CC   of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC   important in breast cancer such as the adenomatous polyposis coli gene
CC   (APC) and the Drosophila discs-large gene (DIg), which are components of
CC   cell-cell junctions. The methods are useful in cardiovascular
CC   applications, neurobiology applications, bone biology applications, skin
CC   biology applications, cosmeceutical applications, endocrinology
CC   applications, infectious disease applications, drug toxicities and drug
CC   resistance applications, immunobiology, inflammation, and allergic
CC   response applications, and biotechnology applications. The peptide
CC   library can easily be monitored, both for its presence within cells and
CC   its quantity. The expression of structurally biased libraries generate
CC   elevated cellular concentration of peptides having a given structural
CC   bias and thus increase the hit rate for targets that bind such
CC   structures. This is the amino acid sequence of a scaffold protein used in
CC   peptide libraries ot hold the library peptide in a conformationally
CC   restricted form.
XX
SQ   Sequence 104 AA;

  Query Match             61.0%;  Score 64;  DB 7;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.45;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|||| |||  || ||| ||||:|
Db          9 AAAEAAAKAA--AAAAAEAAAKAAA 31


RESULT 11
ADE10633
ID   ADE10633 standard; protein; 104 AA.
XX
AC   ADE10633;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Structurally biased random peptide library related protein seqid 40.
XX
KW   fusion nucleic acid library; scaffold protein; bioactive peptide;
KW   phenotype change; cell morphology; cell growth; cell viability;
KW   cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW   loss of cell division; decreased cell growth; brca-1; brca-2;
KW   tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW   Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW   skin biology; cosmeceutical; endocrinology; infectious disease;
KW   drug toxicity; drug resistance; inflammation; allergic response.
XX
OS   Synthetic.
XX
PN   US2003143562-A1.
XX
PD   31-JUL-2003.
XX
PF   20-JUN-2002; 2002US-00177725.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
```

PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2003-829786/77.
XX
PT Novel library of fusion nucleic acids each of which has fused first and
PT second nucleic acids encoding scaffold protein and library peptide having
PT alpha helical biasing sequence, respectively, useful in screening
PT methods.
XX
PS Example 6; SEQ ID NO 40; 110pp; English.
XX
CC The invention describes a library (I) of fusion nucleic acids, where each
CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC library peptide sequence comprising an alpha helical biasing sequence;
CC where N1 is fused to N2. Disclosed is a method for screening bioactive
CC peptides conferring a change in specific phenotype such as cell
CC morphology, cell growth, cell viability, adhesion to substrates or other
CC cells, and cellular density; changes in the expression of one or more
CC RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC peptide identified by above mentioned method is used to generate more
CC candidate peptides and to identify target molecules, i.e., the molecules
CC with which the bioactive peptide interacts. The peptide(s) can be
CC combined with other pharmacologic activators to study the epistatic
CC relationships of signal transduction pathways in question. The disclosed
CC method is also useful in cancer applications. Random libraries can be
CC introduced into any tumour cell (primary or cultured), and peptides
CC identified which by themselves induce apoptosis, cell death, loss of cell
CC division or decreased cell growth. The method is also useful for
CC screening of bioactive peptides which restore the constitutive function
CC of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC important in breast cancer such as the adenomatous polyposis coli gene
CC (APC) and the Drosophila discs-large gene (DIg), which are components of
CC cell-cell junctions. The methods are useful in cardiovascular
CC applications, neurobiology applications, bone biology applications, skin
CC biology applications, cosmeceutical applications, endocrinology
CC applications, infectious disease applications, drug toxicities and drug
CC resistance applications, immunobiology, inflammation, and allergic
CC response applications, and biotechnology applications. The peptide
CC library can easily be monitored, both for its presence within cells and
CC its quantity. The expression of structurally biased libraries generate
CC elevated cellular concentration of peptides having a given structural
CC bias and thus increase the hit rate for targets that bind such
CC structures. This is the amino acid sequence of a protein associated with
CC fused nucleic acid and random peptide libraries of the invention.
XX
SQ Sequence 104 AA;

Query Match 61.0%; Score 64; DB 7; Length 104;
Best Local Similarity 72.0%; Pred. No. 0.45;
Matches 18; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

Qy 1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
|:|||| ||| || ||| ||||:|
Db 9 AAAEAAAKAA--AAAAAEAAAKAAA 31

RESULT 12
ADE10632
ID ADE10632 standard; protein; 104 AA.
XX
AC ADE10632;
XX
DT 29-JAN-2004 (first entry)
XX
DE Structurally biased random peptide library related protein seqid 39.
XX
KW fusion nucleic acid library; scaffold protein; bioactive peptide;
KW phenotype change; cell morphology; cell growth; cell viability;
KW cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW loss of cell division; decreased cell growth; brca-1; brca-2;
KW tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW skin biology; cosmeceutical; endocrinology; infectious disease;
KW drug toxicity; drug resistance; inflammation; allergic response.
XX
OS Synthetic.
XX
PN US2003143562-A1.
XX
PD 31-JUL-2003.
XX
PF 20-JUN-2002; 2002US-00177725.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2003-829786/77.
XX
PT Novel library of fusion nucleic acids each of which has fused first and
PT second nucleic acids encoding scaffold protein and library peptide having
PT alpha helical biasing sequence, respectively, useful in screening
PT methods.
XX
PS Example 6; SEQ ID NO 39; 110pp; English.
XX
CC The invention describes a library (I) of fusion nucleic acids, where each
CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC library peptide sequence comprising an alpha helical biasing sequence;
CC where N1 is fused to N2. Disclosed is a method for screening bioactive
CC peptides conferring a change in specific phenotype such as cell
CC morphology, cell growth, cell viability, adhesion to substrates or other
CC cells, and cellular density; changes in the expression of one or more
CC RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC peptide identified by above mentioned method is used to generate more
CC candidate peptides and to identify target molecules, i.e., the molecules
CC with which the bioactive peptide interacts. The peptide(s) can be
CC combined with other pharmacologic activators to study the epistatic
CC relationships of signal transduction pathways in question. The disclosed
CC method is also useful in cancer applications. Random libraries can be
CC introduced into any tumour cell (primary or cultured), and peptides
CC identified which by themselves induce apoptosis, cell death, loss of cell
CC division or decreased cell growth. The method is also useful for
CC screening of bioactive peptides which restore the constitutive function
CC of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC important in breast cancer such as the adenomatous polyposis coli gene
CC (APC) and the Drosophila discs-large gene (DIg), which are components of
CC cell-cell junctions. The methods are useful in cardiovascular
CC applications, neurobiology applications, bone biology applications, skin
CC biology applications, cosmeceutical applications, endocrinology
CC applications, infectious disease applications, drug toxicities and drug
CC resistance applications, immunobiology, inflammation, and allergic
CC response applications, and biotechnology applications. The peptide
CC library can easily be monitored, both for its presence within cells and
CC its quantity. The expression of structurally biased libraries generate
CC elevated cellular concentration of peptides having a given structural
CC bias and thus increase the hit rate for targets that bind such
CC structures. This is the amino acid sequence of a protein associated with
CC fused nucleic acid and random peptide libraries of the invention.
XX
SQ Sequence 104 AA;

Query Match 61.0%; Score 64; DB 7; Length 104;
Best Local Similarity 72.0%; Pred. No. 0.45;
Matches 18; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

Qy 1 AXAEAAEKAAKYAAEAAEKAAKAXA 25

```
              |:|||| |||  || ||| ||||:|
Db          9 AAAEAAAKAA--AAAAAEAAAKAAA 31


RESULT 13
ADK15652
ID   ADK15652 standard; peptide; 104 AA.
XX
AC   ADK15652;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Nucleating sequence-containing library fusion protein #34.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2004-033956/03.
XX
PT   Library of fusion polypeptides in which each polypeptides comprises
PT   scaffold protein and library peptide having alpha helical biasing
PT   sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS   Example 6; SEQ ID NO 40; 110pp; English.
XX
CC   The invention comprises a library of fusion nucleic acids, where each
CC   encoded protein contains a scaffold protein (e.g. a green fluorescent
CC   protein - GFP) and a library peptide sequence comprising an alpha helical
CC   biasing sequence, or a scaffold protein, a library peptide and a
CC   nucleating sequence. The library of the invention is useful for screening
CC   bioactive peptides conferring a particular phenotype. The present amino
CC   acid sequence represents a library protein containing a nucleating
CC   sequence.
XX
SQ   Sequence 104 AA;

  Query Match             61.0%;  Score 64;  DB 8;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.45;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|||| |||  || ||| ||||:|
Db          9 AAAEAAAKAA--AAAAAEAAAKAAA 31


RESULT 14
ADK15701
ID   ADK15701 standard; peptide; 104 AA.
XX
AC   ADK15701;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Library fusion protein-related scaffold protein #33.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2004-033956/03.
XX
PT   Library of fusion polypeptides in which each polypeptides comprises
PT   scaffold protein and library peptide having alpha helical biasing
PT   sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS   Disclosure; SEQ ID NO 89; 110pp; English.
XX
CC   The invention comprises a library of fusion nucleic acids, where each
CC   encoded protein contains a scaffold protein (e.g. a green fluorescent
CC   protein - GFP) and a library peptide sequence comprising an alpha helical
CC   biasing sequence, or a scaffold protein, a library peptide and a
CC   nucleating sequence. The library of the invention is useful for screening
CC   bioactive peptides conferring a particular phenotype. The present amino
CC   acid sequence represents a scaffold protein.
XX
SQ   Sequence 104 AA;

  Query Match             61.0%;  Score 64;  DB 8;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.45;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|||| |||  || ||| ||||:|
Db          9 AAAEAAAKAA--AAAAAEAAAKAAA 31


RESULT 15
ADK15702
ID   ADK15702 standard; peptide; 104 AA.
XX
AC   ADK15702;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Library fusion protein-related scaffold protein #34.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
```

PR 08-OCT-1999; 99US-00415765.
PR 20-JUN-2002; 2002US-00177725.
XX
PA (ANDE/) ANDERSON D.
PA (PEEL/) PEELLE B R.
PA (BOGE/) BOGENBERGER J M.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2004-033956/03.
XX
PT Library of fusion polypeptides in which each polypeptides comprises
PT scaffold protein and library peptide having alpha helical biasing
PT sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS Disclosure; SEQ ID NO 90; 110pp; English.
XX
CC The invention comprises a library of fusion nucleic acids, where each
CC encoded protein contains a scaffold protein (e.g. a green fluorescent
CC protein - GFP) and a library peptide sequence comprising an alpha helical
CC biasing sequence, or a scaffold protein, a library peptide and a
CC nucleating sequence. The library of the invention is useful for screening
CC bioactive peptides conferring a particular phenotype. The present amino
CC acid sequence represents a scaffold protein.
XX
SQ Sequence 104 AA;

Query Match 61.0%; Score 64; DB 8; Length 104;
Best Local Similarity 72.0%; Pred. No. 0.45;
Matches 18; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

Qy 1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
|:|||| ||| || ||| ||||:|
Db 9 AAAEAAAKAA--AAAAAEAAAKAAA 31

RESULT 16
ADK15651
ID ADK15651 standard; peptide; 104 AA.
XX
AC ADK15651;
XX
DT 06-MAY-2004 (first entry)
XX
DE Nucleating sequence-containing library fusion protein #33.
XX
KW fusion nucleic acid library; fusion protein library; scaffold protein;
KW green fluorescent protein; GFP; alpha helical biasing sequence;
KW nucleating sequence; screening.
XX
OS Synthetic.
XX
PN US2003224412-A1.
XX
PD 04-DEC-2003.
XX
PF 18-MAR-2003; 2003US-00393449.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
PR 20-JUN-2002; 2002US-00177725.
XX
PA (ANDE/) ANDERSON D.
PA (PEEL/) PEELLE B R.
PA (BOGE/) BOGENBERGER J M.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2004-033956/03.
XX
PT Library of fusion polypeptides in which each polypeptides comprises
PT scaffold protein and library peptide having alpha helical biasing
PT sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS Example 6; SEQ ID NO 39; 110pp; English.
XX
CC The invention comprises a library of fusion nucleic acids, where each
CC encoded protein contains a scaffold protein (e.g. a green fluorescent
CC protein - GFP) and a library peptide sequence comprising an alpha helical
CC biasing sequence, or a scaffold protein, a library peptide and a
CC nucleating sequence. The library of the invention is useful for screening
CC bioactive peptides conferring a particular phenotype. The present amino
CC acid sequence represents a library protein containing a nucleating
CC sequence.
XX
SQ Sequence 104 AA;

Query Match 61.0%; Score 64; DB 8; Length 104;
Best Local Similarity 72.0%; Pred. No. 0.45;
Matches 18; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

Qy 1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
|:|||| ||| || ||| ||||:|
Db 9 AAAEAAAKAA--AAAAAEAAAKAAA 31

RESULT 17
ADE10684
ID ADE10684 standard; protein; 104 AA.
XX
AC ADE10684;
XX
DT 29-JAN-2004 (first entry)
XX
DE Structurally biased random peptide library scaffold protein seqid 91.
XX
KW fusion nucleic acid library; scaffold protein; bioactive peptide;
KW phenotype change; cell morphology; cell growth; cell viability;
KW cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW loss of cell division; decreased cell growth; brca-1; brca-2;
KW tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW skin biology; cosmeceutical; endocrinology; infectious disease;
KW drug toxicity; drug resistance; inflammation; allergic response;
KW scaffold protein.
XX
OS Synthetic.
XX
PN US2003143562-A1.
XX
PD 31-JUL-2003.
XX
PF 20-JUN-2002; 2002US-00177725.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2003-829786/77.
XX
PT Novel library of fusion nucleic acids each of which has fused first and
PT second nucleic acids encoding scaffold protein and library peptide having
PT alpha helical biasing sequence, respectively, useful in screening
PT methods.
XX
PS Disclosure; SEQ ID NO 91; 110pp; English.
XX
CC The invention describes a library (I) of fusion nucleic acids, where each
CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC library peptide sequence comprising an alpha helical biasing sequence;
CC where N1 is fused to N2. Disclosed is a method for screening bioactive
CC peptides conferring a change in specific phenotype such as cell

CC morphology, cell growth, cell viability, adhesion to substrates or other
CC cells, and cellular density; changes in the expression of one or more
CC RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC peptide identified by above mentioned method is used to generate more
CC candidate peptides and to identify target molecules, i.e., the molecules
CC with which the bioactive peptide interacts. The peptide(s) can be
CC combined with other pharmacologic activators to study the epistatic
CC relationships of signal transduction pathways in question. The disclosed
CC method is also useful in cancer applications. Random libraries can be
CC introduced into any tumour cell (primary or cultured), and peptides
CC identified which by themselves induce apoptosis, cell death, loss of cell
CC division or decreased cell growth. The method is also useful for
CC screening of bioactive peptides which restore the constitutive function
CC of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC important in breast cancer such as the adenomatous polyposis coli gene
CC (APC) and the Drosophila discs-large gene (DIg), which are components of
CC cell-cell junctions. The methods are useful in cardiovascular
CC applications, neurobiology applications, bone biology applications, skin
CC biology applications, cosmeceutical applications, endocrinology
CC applications, infectious disease applications, drug toxicities and drug
CC resistance applications, immunobiology, inflammation, and allergic
CC response applications, and biotechnology applications. The peptide
CC library can easily be monitored, both for its presence within cells and
CC its quantity. The expression of structurally biased libraries generate
CC elevated cellular concentration of peptides having a given structural
CC bias and thus increase the hit rate for targets that bind such
CC structures. This is the amino acid sequence of a scaffold protein used in
CC peptide libraries ot hold the library peptide in a conformationally
CC restricted form.
XX
SQ Sequence 104 AA;

Query Match 60.5%; Score 63.5; DB 7; Length 104;
Best Local Similarity 72.0%; Pred. No. 0.52;
Matches 18; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

```
Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||| ||| || ||| ||||:|
Db          6 AAAAAAEAAAK-AAAAAEAAAKAAA 29
```

RESULT 18
ADE10634
ID ADE10634 standard; protein; 104 AA.
XX
AC ADE10634;
XX
DT 29-JAN-2004 (first entry)
XX
DE Structurally biased random peptide library related protein seqid 41.
XX
KW fusion nucleic acid library; scaffold protein; bioactive peptide;
KW phenotype change; cell morphology; cell growth; cell viability;
KW cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW loss of cell division; decreased cell growth; brca-1; brca-2;
KW tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW skin biology; cosmeceutical; endocrinology; infectious disease;
KW drug toxicity; drug resistance; inflammation; allergic response.
XX
OS Synthetic.
XX
PN US2003143562-A1.
XX
PD 31-JUL-2003.
XX
PF 20-JUN-2002; 2002US-00177725.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2003-829786/77.
XX
PT Novel library of fusion nucleic acids each of which has fused first and
PT second nucleic acids encoding scaffold protein and library peptide having
PT alpha helical biasing sequence, respectively, useful in screening
PT methods.
XX
PS Example 6; SEQ ID NO 41; 110pp; English.
XX
CC The invention describes a library (I) of fusion nucleic acids, where each
CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC library peptide sequence comprising an alpha helical biasing sequence;
CC where N1 is fused to N2. Disclosed is a method for screening bioactive
CC peptides conferring a change in specific phenotype such as cell
CC morphology, cell growth, cell viability, adhesion to substrates or other
CC cells, and cellular density; changes in the expression of one or more
CC RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC peptide identified by above mentioned method is used to generate more
CC candidate peptides and to identify target molecules, i.e., the molecules
CC with which the bioactive peptide interacts. The peptide(s) can be
CC combined with other pharmacologic activators to study the epistatic
CC relationships of signal transduction pathways in question. The disclosed
CC method is also useful in cancer applications. Random libraries can be
CC introduced into any tumour cell (primary or cultured), and peptides
CC identified which by themselves induce apoptosis, cell death, loss of cell
CC division or decreased cell growth. The method is also useful for
CC screening of bioactive peptides which restore the constitutive function
CC of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC important in breast cancer such as the adenomatous polyposis coli gene
CC (APC) and the Drosophila discs-large gene (DIg), which are components of
CC cell-cell junctions. The methods are useful in cardiovascular
CC applications, neurobiology applications, bone biology applications, skin
CC biology applications, cosmeceutical applications, endocrinology
CC applications, infectious disease applications, drug toxicities and drug
CC resistance applications, immunobiology, inflammation, and allergic
CC response applications, and biotechnology applications. The peptide
CC library can easily be monitored, both for its presence within cells and
CC its quantity. The expression of structurally biased libraries generate
CC elevated cellular concentration of peptides having a given structural
CC bias and thus increase the hit rate for targets that bind such
CC structures. This is the amino acid sequence of a protein associated with
CC fused nucleic acid and random peptide libraries of the invention.
XX
SQ Sequence 104 AA;

Query Match 60.5%; Score 63.5; DB 7; Length 104;
Best Local Similarity 72.0%; Pred. No. 0.52;
Matches 18; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

```
Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||| ||| || ||| ||||:|
Db          6 AAAAAAEAAAK-AAAAAEAAAKAAA 29
```

RESULT 19
ADK15703
ID ADK15703 standard; peptide; 104 AA.
XX
AC ADK15703;
XX
DT 06-MAY-2004 (first entry)
XX
DE Library fusion protein-related scaffold protein #35.
XX
KW fusion nucleic acid library; fusion protein library; scaffold protein;
KW green fluorescent protein; GFP; alpha helical biasing sequence;

KW nucleating sequence; screening.
XX
OS Synthetic.
XX
PN US2003224412-A1.
XX
PD 04-DEC-2003.
XX
PF 18-MAR-2003; 2003US-00393449.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
PR 20-JUN-2002; 2002US-00177725.
XX
PA (ANDE/) ANDERSON D.
PA (PEEL/) PEELLE B R.
PA (BOGE/) BOGENBERGER J M.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2004-033956/03.
XX
PT Library of fusion polypeptides in which each polypeptides comprises
PT scaffold protein and library peptide having alpha helical biasing
PT sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS Disclosure; SEQ ID NO 91; 110pp; English.
XX
CC The invention comprises a library of fusion nucleic acids, where each
CC encoded protein contains a scaffold protein (e.g. a green fluorescent
CC protein - GFP) and a library peptide sequence comprising an alpha helical
CC biasing sequence, or a scaffold protein, a library peptide and a
CC nucleating sequence. The library of the invention is useful for screening
CC bioactive peptides conferring a particular phenotype. The present amino
CC acid sequence represents a scaffold protein.
XX
SQ Sequence 104 AA;

Query Match 60.5%; Score 63.5; DB 8; Length 104;
Best Local Similarity 72.0%; Pred. No. 0.52;
Matches 18; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
|:| ||| ||| || ||| ||||:|
Db 6 AAAAAAEAAAK-AAAAAEAAAKAAA 29

RESULT 20
ADK15653
ID ADK15653 standard; peptide; 104 AA.
XX
AC ADK15653;
XX
DT 06-MAY-2004 (first entry)
XX
DE Nucleating sequence-containing library fusion protein #35.
XX
KW fusion nucleic acid library; fusion protein library; scaffold protein;
KW green fluorescent protein; GFP; alpha helical biasing sequence;
KW nucleating sequence; screening.
XX
OS Synthetic.
XX
PN US2003224412-A1.
XX
PD 04-DEC-2003.
XX
PF 18-MAR-2003; 2003US-00393449.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
PR 20-JUN-2002; 2002US-00177725.
XX
PA (ANDE/) ANDERSON D.
PA (PEEL/) PEELLE B R.
PA (BOGE/) BOGENBERGER J M.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2004-033956/03.
XX
PT Library of fusion polypeptides in which each polypeptides comprises
PT scaffold protein and library peptide having alpha helical biasing
PT sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS Example 6; SEQ ID NO 41; 110pp; English.
XX
CC The invention comprises a library of fusion nucleic acids, where each
CC encoded protein contains a scaffold protein (e.g. a green fluorescent
CC protein - GFP) and a library peptide sequence comprising an alpha helical
CC biasing sequence, or a scaffold protein, a library peptide and a
CC nucleating sequence. The library of the invention is useful for screening
CC bioactive peptides conferring a particular phenotype. The present amino
CC acid sequence represents a library protein containing a nucleating
CC sequence.
XX
SQ Sequence 104 AA;

Query Match 60.5%; Score 63.5; DB 8; Length 104;
Best Local Similarity 72.0%; Pred. No. 0.52;
Matches 18; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
|:| ||| ||| || ||| ||||:|
Db 6 AAAAAAEAAAK-AAAAAEAAAKAAA 29

RESULT 21
ADE10698
ID ADE10698 standard; protein; 59 AA.
XX
AC ADE10698;
XX
DT 29-JAN-2004 (first entry)
XX
DE Structurally biased random peptide library scaffold protein seqid 105.
XX
KW fusion nucleic acid library; scaffold protein; bioactive peptide;
KW phenotype change; cell morphology; cell growth; cell viability;
KW cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW loss of cell division; decreased cell growth; brca-1; brca-2;
KW tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW skin biology; cosmeceutical; endocrinology; infectious disease;
KW drug toxicity; drug resistance; inflammation; allergic response;
KW scaffold protein.
XX
OS Synthetic.
XX
PN US2003143562-A1.
XX
PD 31-JUL-2003.
XX
PF 20-JUN-2002; 2002US-00177725.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2003-829786/77.
XX
PT Novel library of fusion nucleic acids each of which has fused first and
PT second nucleic acids encoding scaffold protein and library peptide having

PT alpha helical biasing sequence, respectively, useful in screening
PT methods.
XX
PS Disclosure; SEQ ID NO 105; 110pp; English.
XX
CC The invention describes a library (I) of fusion nucleic acids, where each
CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC library peptide sequence comprising an alpha helical biasing sequence;
CC where N1 is fused to N2. Disclosed is a method for screening bioactive
CC peptides conferring a change in specific phenotype such as cell
CC morphology, cell growth, cell viability, adhesion to substrates or other
CC cells, and cellular density; changes in the expression of one or more
CC RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC peptide identified by above mentioned method is used to generate more
CC candidate peptides and to identify target molecules, i.e., the molecules
CC with which the bioactive peptide interacts. The peptide(s) can be
CC combined with other pharmacologic activators to study the epistatic
CC relationships of signal transduction pathways in question. The disclosed
CC method is also useful in cancer applications. Random libraries can be
CC introduced into any tumour cell (primary or cultured), and peptides
CC identified which by themselves induce apoptosis, cell death, loss of cell
CC division or decreased cell growth. The method is also useful for
CC screening of bioactive peptides which restore the constitutive function
CC of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC important in breast cancer such as the adenomatous polyposis coli gene
CC (APC) and the Drosophila discs-large gene (DIg), which are components of
CC cell-cell junctions. The methods are useful in cardiovascular
CC applications, neurobiology applications, bone biology applications, skin
CC biology applications, cosmeceutical applications, endocrinology
CC applications, infectious disease applications, drug toxicities and drug
CC resistance applications, immunobiology, inflammation, and allergic
CC response applications, and biotechnology applications. The peptide
CC library can easily be monitored, both for its presence within cells and
CC its quantity. The expression of structurally biased libraries generate
CC elevated cellular concentration of peptides having a given structural
CC bias and thus increase the hit rate for targets that bind such
CC structures. This is the amino acid sequence of a scaffold protein used in
CC peptide libraries ot hold the library peptide in a conformationally
CC restricted form.
XX
SQ Sequence 59 AA;

```
  Query Match             60.0%;  Score 63;  DB 7;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 0.32;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27
```

RESULT 22
ADE10648
ID ADE10648 standard; protein; 59 AA.
XX
AC ADE10648;
XX
DT 29-JAN-2004 (first entry)
XX
DE Structurally biased random peptide library related protein seqid 55.
XX
KW fusion nucleic acid library; scaffold protein; bioactive peptide;
KW phenotype change; cell morphology; cell growth; cell viability;
KW cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW loss of cell division; decreased cell growth; brca-1; brca-2;
KW tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW skin biology; cosmeceutical; endocrinology; infectious disease;
KW drug toxicity; drug resistance; inflammation; allergic response.
XX
OS Synthetic.
XX
PN US2003143562-A1.
XX
PD 31-JUL-2003.
XX
PF 20-JUN-2002; 2002US-00177725.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2003-829786/77.
XX
PT Novel library of fusion nucleic acids each of which has fused first and
PT second nucleic acids encoding scaffold protein and library peptide having
PT alpha helical biasing sequence, respectively, useful in screening
PT methods.
XX
PS Example 6; SEQ ID NO 55; 110pp; English.
XX
CC The invention describes a library (I) of fusion nucleic acids, where each
CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC library peptide sequence comprising an alpha helical biasing sequence;
CC where N1 is fused to N2. Disclosed is a method for screening bioactive
CC peptides conferring a change in specific phenotype such as cell
CC morphology, cell growth, cell viability, adhesion to substrates or other
CC cells, and cellular density; changes in the expression of one or more
CC RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC peptide identified by above mentioned method is used to generate more
CC candidate peptides and to identify target molecules, i.e., the molecules
CC with which the bioactive peptide interacts. The peptide(s) can be
CC combined with other pharmacologic activators to study the epistatic
CC relationships of signal transduction pathways in question. The disclosed
CC method is also useful in cancer applications. Random libraries can be
CC introduced into any tumour cell (primary or cultured), and peptides
CC identified which by themselves induce apoptosis, cell death, loss of cell
CC division or decreased cell growth. The method is also useful for
CC screening of bioactive peptides which restore the constitutive function
CC of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC important in breast cancer such as the adenomatous polyposis coli gene
CC (APC) and the Drosophila discs-large gene (DIg), which are components of
CC cell-cell junctions. The methods are useful in cardiovascular
CC applications, neurobiology applications, bone biology applications, skin
CC biology applications, cosmeceutical applications, endocrinology
CC applications, infectious disease applications, drug toxicities and drug
CC resistance applications, immunobiology, inflammation, and allergic
CC response applications, and biotechnology applications. The peptide
CC library can easily be monitored, both for its presence within cells and
CC its quantity. The expression of structurally biased libraries generate
CC elevated cellular concentration of peptides having a given structural
CC bias and thus increase the hit rate for targets that bind such
CC structures. This is the amino acid sequence of a protein associated with
CC fused nucleic acid and random peptide libraries of the invention.
XX
SQ Sequence 59 AA;

```
  Query Match             60.0%;  Score 63;  DB 7;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 0.32;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27
```

RESULT 23

```
ADK15717
ID   ADK15717 standard; peptide; 59 AA.
XX
AC   ADK15717;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Library fusion protein-related scaffold protein #49.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2004-033956/03.
XX
PT   Library of fusion polypeptides in which each polypeptides comprises
PT   scaffold protein and library peptide having alpha helical biasing
PT   sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS   Disclosure; SEQ ID NO 105; 110pp; English.
XX
CC   The invention comprises a library of fusion nucleic acids, where each
CC   encoded protein contains a scaffold protein (e.g. a green fluorescent
CC   protein - GFP) and a library peptide sequence comprising an alpha helical
CC   biasing sequence, or a scaffold protein, a library peptide and a
CC   nucleating sequence. The library of the invention is useful for screening
CC   bioactive peptides conferring a particular phenotype. The present amino
CC   acid sequence represents a scaffold protein.
XX
SQ   Sequence 59 AA;

  Query Match             60.0%;  Score 63;  DB 8;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 0.32;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 24
ADK15667
ID   ADK15667 standard; peptide; 59 AA.
XX
AC   ADK15667;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Nucleating sequence-containing library fusion protein #49.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2004-033956/03.
XX
PT   Library of fusion polypeptides in which each polypeptides comprises
PT   scaffold protein and library peptide having alpha helical biasing
PT   sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS   Example 6; SEQ ID NO 55; 110pp; English.
XX
CC   The invention comprises a library of fusion nucleic acids, where each
CC   encoded protein contains a scaffold protein (e.g. a green fluorescent
CC   protein - GFP) and a library peptide sequence comprising an alpha helical
CC   biasing sequence, or a scaffold protein, a library peptide and a
CC   nucleating sequence. The library of the invention is useful for screening
CC   bioactive peptides conferring a particular phenotype. The present amino
CC   acid sequence represents a library protein containing a nucleating
CC   sequence.
XX
SQ   Sequence 59 AA;

  Query Match             60.0%;  Score 63;  DB 8;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 0.32;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 25
ADE10697
ID   ADE10697 standard; protein; 67 AA.
XX
AC   ADE10697;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Structurally biased random peptide library scaffold protein seqid 104.
XX
KW   fusion nucleic acid library; scaffold protein; bioactive peptide;
KW   phenotype change; cell morphology; cell growth; cell viability;
KW   cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW   loss of cell division; decreased cell growth; brca-1; brca-2;
KW   tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW   Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW   skin biology; cosmeceutical; endocrinology; infectious disease;
KW   drug toxicity; drug resistance; inflammation; allergic response;
KW   scaffold protein.
XX
OS   Synthetic.
XX
PN   US2003143562-A1.
XX
PD   31-JUL-2003.
XX
PF   20-JUN-2002; 2002US-00177725.
XX
```

PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2003-829786/77.
XX
PT Novel library of fusion nucleic acids each of which has fused first and
PT second nucleic acids encoding scaffold protein and library peptide having
PT alpha helical biasing sequence, respectively, useful in screening
PT methods.
XX
PS Disclosure; SEQ ID NO 104; 110pp; English.
XX
CC The invention describes a library (I) of fusion nucleic acids, where each
CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC library peptide sequence comprising an alpha helical biasing sequence;
CC where N1 is fused to N2. Disclosed is a method for screening bioactive
CC peptides conferring a change in specific phenotype such as cell
CC morphology, cell growth, cell viability, adhesion to substrates or other
CC cells, and cellular density; changes in the expression of one or more
CC RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC peptide identified by above mentioned method is used to generate more
CC candidate peptides and to identify target molecules, i.e., the molecules
CC with which the bioactive peptide interacts. The peptide(s) can be
CC combined with other pharmacologic activators to study the epistatic
CC relationships of signal transduction pathways in question. The disclosed
CC method is also useful in cancer applications. Random libraries can be
CC introduced into any tumour cell (primary or cultured), and peptides
CC identified which by themselves induce apoptosis, cell death, loss of cell
CC division or decreased cell growth. The method is also useful for
CC screening of bioactive peptides which restore the constitutive function
CC of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC important in breast cancer such as the adenomatous polyposis coli gene
CC (APC) and the Drosophila discs-large gene (DIg), which are components of
CC cell-cell junctions. The methods are useful in cardiovascular
CC applications, neurobiology applications, bone biology applications, skin
CC biology applications, cosmeceutical applications, endocrinology
CC applications, infectious disease applications, drug toxicities and drug
CC resistance applications, immunobiology, inflammation, and allergic
CC response applications, and biotechnology applications. The peptide
CC library can easily be monitored, both for its presence within cells and
CC its quantity. The expression of structurally biased libraries generate
CC elevated cellular concentration of peptides having a given structural
CC bias and thus increase the hit rate for targets that bind such
CC structures. This is the amino acid sequence of a scaffold protein used in
CC peptide libraries ot hold the library peptide in a conformationally
CC restricted form.
XX
SQ Sequence 67 AA;

Query Match 60.0%; Score 63; DB 7; Length 67;
Best Local Similarity 66.7%; Pred. No. 0.37;
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 XAEAAEKAAKYAAEAAEKAAKAXA 25
:| ||| ||| ||||| ||| :|
Db 4 DAAAAEAAAKAAAEAAAKAAAEAA 27

RESULT 26
ADE10647
ID ADE10647 standard; protein; 67 AA.
XX
AC ADE10647;
XX
DT 29-JAN-2004 (first entry)
XX
DE Structurally biased random peptide library related protein seqid 54.
XX
KW fusion nucleic acid library; scaffold protein; bioactive peptide;
KW phenotype change; cell morphology; cell growth; cell viability;
KW cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW loss of cell division; decreased cell growth; brca-1; brca-2;
KW tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW skin biology; cosmeceutical; endocrinology; infectious disease;
KW drug toxicity; drug resistance; inflammation; allergic response.
XX
OS Synthetic.
XX
PN US2003143562-A1.
XX
PD 31-JUL-2003.
XX
PF 20-JUN-2002; 2002US-00177725.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2003-829786/77.
XX
PT Novel library of fusion nucleic acids each of which has fused first and
PT second nucleic acids encoding scaffold protein and library peptide having
PT alpha helical biasing sequence, respectively, useful in screening
PT methods.
XX
PS Example 6; SEQ ID NO 54; 110pp; English.
XX
CC The invention describes a library (I) of fusion nucleic acids, where each
CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC library peptide sequence comprising an alpha helical biasing sequence;
CC where N1 is fused to N2. Disclosed is a method for screening bioactive
CC peptides conferring a change in specific phenotype such as cell
CC morphology, cell growth, cell viability, adhesion to substrates or other
CC cells, and cellular density; changes in the expression of one or more
CC RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC peptide identified by above mentioned method is used to generate more
CC candidate peptides and to identify target molecules, i.e., the molecules
CC with which the bioactive peptide interacts. The peptide(s) can be
CC combined with other pharmacologic activators to study the epistatic
CC relationships of signal transduction pathways in question. The disclosed
CC method is also useful in cancer applications. Random libraries can be
CC introduced into any tumour cell (primary or cultured), and peptides
CC identified which by themselves induce apoptosis, cell death, loss of cell
CC division or decreased cell growth. The method is also useful for
CC screening of bioactive peptides which restore the constitutive function
CC of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC important in breast cancer such as the adenomatous polyposis coli gene
CC (APC) and the Drosophila discs-large gene (DIg), which are components of
CC cell-cell junctions. The methods are useful in cardiovascular
CC applications, neurobiology applications, bone biology applications, skin
CC biology applications, cosmeceutical applications, endocrinology
CC applications, infectious disease applications, drug toxicities and drug
CC resistance applications, immunobiology, inflammation, and allergic
CC response applications, and biotechnology applications. The peptide
CC library can easily be monitored, both for its presence within cells and
CC its quantity. The expression of structurally biased libraries generate
CC elevated cellular concentration of peptides having a given structural
CC bias and thus increase the hit rate for targets that bind such
CC structures. This is the amino acid sequence of a protein associated with
CC fused nucleic acid and random peptide libraries of the invention.
XX
SQ Sequence 67 AA;

```
  Query Match             60.0%;  Score 63;  DB 7;  Length 67;
  Best Local Similarity   66.7%;  Pred. No. 0.37;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 27
ADK15666
ID   ADK15666 standard; peptide; 67 AA.
XX
AC   ADK15666;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Nucleating sequence-containing library fusion protein #48.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2004-033956/03.
XX
PT   Library of fusion polypeptides in which each polypeptides comprises
PT   scaffold protein and library peptide having alpha helical biasing
PT   sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS   Example 6; SEQ ID NO 54; 110pp; English.
XX
CC   The invention comprises a library of fusion nucleic acids, where each
CC   encoded protein contains a scaffold protein (e.g. a green fluorescent
CC   protein - GFP) and a library peptide sequence comprising an alpha helical
CC   biasing sequence, or a scaffold protein, a library peptide and a
CC   nucleating sequence. The library of the invention is useful for screening
CC   bioactive peptides conferring a particular phenotype. The present amino
CC   acid sequence represents a library protein containing a nucleating
CC   sequence.
XX
SQ   Sequence 67 AA;

  Query Match             60.0%;  Score 63;  DB 8;  Length 67;
  Best Local Similarity   66.7%;  Pred. No. 0.37;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 28
ADK15716
ID   ADK15716 standard; peptide; 67 AA.
XX
AC   ADK15716;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Library fusion protein-related scaffold protein #48.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2004-033956/03.
XX
PT   Library of fusion polypeptides in which each polypeptides comprises
PT   scaffold protein and library peptide having alpha helical biasing
PT   sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS   Disclosure; SEQ ID NO 104; 110pp; English.
XX
CC   The invention comprises a library of fusion nucleic acids, where each
CC   encoded protein contains a scaffold protein (e.g. a green fluorescent
CC   protein - GFP) and a library peptide sequence comprising an alpha helical
CC   biasing sequence, or a scaffold protein, a library peptide and a
CC   nucleating sequence. The library of the invention is useful for screening
CC   bioactive peptides conferring a particular phenotype. The present amino
CC   acid sequence represents a scaffold protein.
XX
SQ   Sequence 67 AA;

  Query Match             60.0%;  Score 63;  DB 8;  Length 67;
  Best Local Similarity   66.7%;  Pred. No. 0.37;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 29
ADE10696
ID   ADE10696 standard; protein; 75 AA.
XX
AC   ADE10696;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Structurally biased random peptide library scaffold protein seqid 103.
XX
KW   fusion nucleic acid library; scaffold protein; bioactive peptide;
KW   phenotype change; cell morphology; cell growth; cell viability;
KW   cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW   loss of cell division; decreased cell growth; brca-1; brca-2;
KW   tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW   Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW   skin biology; cosmeceutical; endocrinology; infectious disease;
```

KW drug toxicity; drug resistance; inflammation; allergic response;
KW scaffold protein.
XX
OS Synthetic.
XX
PN US2003143562-A1.
XX
PD 31-JUL-2003.
XX
PF 20-JUN-2002; 2002US-00177725.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2003-829786/77.
XX
PT Novel library of fusion nucleic acids each of which has fused first and
PT second nucleic acids encoding scaffold protein and library peptide having
PT alpha helical biasing sequence, respectively, useful in screening
PT methods.
XX
PS Disclosure; SEQ ID NO 103; 110pp; English.
XX
CC The invention describes a library (I) of fusion nucleic acids, where each
CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC library peptide sequence comprising an alpha helical biasing sequence;
CC where N1 is fused to N2. Disclosed is a method for screening bioactive
CC peptides conferring a change in specific phenotype such as cell
CC morphology, cell growth, cell viability, adhesion to substrates or other
CC cells, and cellular density; changes in the expression of one or more
CC RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC peptide identified by above mentioned method is used to generate more
CC candidate peptides and to identify target molecules, i.e., the molecules
CC with which the bioactive peptide interacts. The peptide(s) can be
CC combined with other pharmacologic activators to study the epistatic
CC relationships of signal transduction pathways in question. The disclosed
CC method is also useful in cancer applications. Random libraries can be
CC introduced into any tumour cell (primary or cultured), and peptides
CC identified which by themselves induce apoptosis, cell death, loss of cell
CC division or decreased cell growth. The method is also useful for
CC screening of bioactive peptides which restore the constitutive function
CC of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC important in breast cancer such as the adenomatous polyposis coli gene
CC (APC) and the Drosophila discs-large gene (DIg), which are components of
CC cell-cell junctions. The methods are useful in cardiovascular
CC applications, neurobiology applications, bone biology applications, skin
CC biology applications, cosmeceutical applications, endocrinology
CC applications, infectious disease applications, drug toxicities and drug
CC resistance applications, immunobiology, inflammation, and allergic
CC response applications, and biotechnology applications. The peptide
CC library can easily be monitored, both for its presence within cells and
CC its quantity. The expression of structurally biased libraries generate
CC elevated cellular concentration of peptides having a given structural
CC bias and thus increase the hit rate for targets that bind such
CC structures. This is the amino acid sequence of a scaffold protein used in
CC peptide libraries ot hold the library peptide in a conformationally
CC restricted form.
XX
SQ Sequence 75 AA;

Query Match 60.0%; Score 63; DB 7; Length 75;
Best Local Similarity 66.7%; Pred. No. 0.42;
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 XAEAAEKAAKYAAEAAEKAAKAXA 25
:| ||| ||| ||||| ||| :|

Db 4 DAAAAEAAAKAAAEAAAKAAAEAA 27

RESULT 30
ADE10646
ID ADE10646 standard; protein; 75 AA.
XX
AC ADE10646;
XX
DT 29-JAN-2004 (first entry)
XX
DE Structurally biased random peptide library related protein seqid 53.
XX
KW fusion nucleic acid library; scaffold protein; bioactive peptide;
KW phenotype change; cell morphology; cell growth; cell viability;
KW cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW loss of cell division; decreased cell growth; brca-1; brca-2;
KW tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW skin biology; cosmeceutical; endocrinology; infectious disease;
KW drug toxicity; drug resistance; inflammation; allergic response.
XX
OS Synthetic.
XX
PN US2003143562-A1.
XX
PD 31-JUL-2003.
XX
PF 20-JUN-2002; 2002US-00177725.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2003-829786/77.
XX
PT Novel library of fusion nucleic acids each of which has fused first and
PT second nucleic acids encoding scaffold protein and library peptide having
PT alpha helical biasing sequence, respectively, useful in screening
PT methods.
XX
PS Example 6; SEQ ID NO 53; 110pp; English.
XX
CC The invention describes a library (I) of fusion nucleic acids, where each
CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC library peptide sequence comprising an alpha helical biasing sequence;
CC where N1 is fused to N2. Disclosed is a method for screening bioactive
CC peptides conferring a change in specific phenotype such as cell
CC morphology, cell growth, cell viability, adhesion to substrates or other
CC cells, and cellular density; changes in the expression of one or more
CC RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC peptide identified by above mentioned method is used to generate more
CC candidate peptides and to identify target molecules, i.e., the molecules
CC with which the bioactive peptide interacts. The peptide(s) can be
CC combined with other pharmacologic activators to study the epistatic
CC relationships of signal transduction pathways in question. The disclosed
CC method is also useful in cancer applications. Random libraries can be
CC introduced into any tumour cell (primary or cultured), and peptides
CC identified which by themselves induce apoptosis, cell death, loss of cell
CC division or decreased cell growth. The method is also useful for
CC screening of bioactive peptides which restore the constitutive function
CC of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC important in breast cancer such as the adenomatous polyposis coli gene
CC (APC) and the Drosophila discs-large gene (DIg), which are components of
CC cell-cell junctions. The methods are useful in cardiovascular
CC applications, neurobiology applications, bone biology applications, skin
CC biology applications, cosmeceutical applications, endocrinology

```
CC   applications, infectious disease applications, drug toxicities and drug
CC   resistance applications, immunobiology, inflammation, and allergic
CC   response applications, and biotechnology applications. The peptide
CC   library can easily be monitored, both for its presence within cells and
CC   its quantity. The expression of structurally biased libraries generate
CC   elevated cellular concentration of peptides having a given structural
CC   bias and thus increase the hit rate for targets that bind such
CC   structures. This is the amino acid sequence of a protein associated with
CC   fused nucleic acid and random peptide libraries of the invention.
XX
SQ   Sequence 75 AA;

  Query Match             60.0%;  Score 63;  DB 7;  Length 75;
  Best Local Similarity   66.7%;  Pred. No. 0.42;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 31
ADK15715
ID   ADK15715 standard; peptide; 75 AA.
XX
AC   ADK15715;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Library fusion protein-related scaffold protein #47.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2004-033956/03.
XX
PT   Library of fusion polypeptides in which each polypeptides comprises
PT   scaffold protein and library peptide having alpha helical biasing
PT   sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS   Disclosure; SEQ ID NO 103; 110pp; English.
XX
CC   The invention comprises a library of fusion nucleic acids, where each
CC   encoded protein contains a scaffold protein (e.g. a green fluorescent
CC   protein - GFP) and a library peptide sequence comprising an alpha helical
CC   biasing sequence, or a scaffold protein, a library peptide and a
CC   nucleating sequence. The library of the invention is useful for screening
CC   bioactive peptides conferring a particular phenotype. The present amino
CC   acid sequence represents a scaffold protein.
XX
SQ   Sequence 75 AA;

  Query Match             60.0%;  Score 63;  DB 8;  Length 75;
  Best Local Similarity   66.7%;  Pred. No. 0.42;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 32
ADK15665
ID   ADK15665 standard; peptide; 75 AA.
XX
AC   ADK15665;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Nucleating sequence-containing library fusion protein #47.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2004-033956/03.
XX
PT   Library of fusion polypeptides in which each polypeptides comprises
PT   scaffold protein and library peptide having alpha helical biasing
PT   sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS   Example 6; SEQ ID NO 53; 110pp; English.
XX
CC   The invention comprises a library of fusion nucleic acids, where each
CC   encoded protein contains a scaffold protein (e.g. a green fluorescent
CC   protein - GFP) and a library peptide sequence comprising an alpha helical
CC   biasing sequence, or a scaffold protein, a library peptide and a
CC   nucleating sequence. The library of the invention is useful for screening
CC   bioactive peptides conferring a particular phenotype. The present amino
CC   acid sequence represents a library protein containing a nucleating
CC   sequence.
XX
SQ   Sequence 75 AA;

  Query Match             60.0%;  Score 63;  DB 8;  Length 75;
  Best Local Similarity   66.7%;  Pred. No. 0.42;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 33
ADE10695
ID   ADE10695 standard; protein; 83 AA.
XX
AC   ADE10695;
XX
```

DT 29-JAN-2004 (first entry)
XX
DE Structurally biased random peptide library scaffold protein seqid 102.
XX
KW fusion nucleic acid library; scaffold protein; bioactive peptide;
KW phenotype change; cell morphology; cell growth; cell viability;
KW cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW loss of cell division; decreased cell growth; brca-1; brca-2;
KW tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW skin biology; cosmeceutical; endocrinology; infectious disease;
KW drug toxicity; drug resistance; inflammation; allergic response;
KW scaffold protein.
XX
OS Synthetic.
XX
PN US2003143562-A1.
XX
PD 31-JUL-2003.
XX
PF 20-JUN-2002; 2002US-00177725.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2003-829786/77.
XX
PT Novel library of fusion nucleic acids each of which has fused first and
PT second nucleic acids encoding scaffold protein and library peptide having
PT alpha helical biasing sequence, respectively, useful in screening
PT methods.
XX
PS Disclosure; SEQ ID NO 102; 110pp; English.
XX
CC The invention describes a library (I) of fusion nucleic acids, where each
CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC library peptide sequence comprising an alpha helical biasing sequence;
CC where N1 is fused to N2. Disclosed is a method for screening bioactive
CC peptides conferring a change in specific phenotype such as cell
CC morphology, cell growth, cell viability, adhesion to substrates or other
CC cells, and cellular density; changes in the expression of one or more
CC RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC peptide identified by above mentioned method is used to generate more
CC candidate peptides and to identify target molecules, i.e., the molecules
CC with which the bioactive peptide interacts. The peptide(s) can be
CC combined with other pharmacologic activators to study the epistatic
CC relationships of signal transduction pathways in question. The disclosed
CC method is also useful in cancer applications. Random libraries can be
CC introduced into any tumour cell (primary or cultured), and peptides
CC identified which by themselves induce apoptosis, cell death, loss of cell
CC division or decreased cell growth. The method is also useful for
CC screening of bioactive peptides which restore the constitutive function
CC of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC important in breast cancer such as the adenomatous polyposis coli gene
CC (APC) and the Drosophila discs-large gene (DIg), which are components of
CC cell-cell junctions. The methods are useful in cardiovascular
CC applications, neurobiology applications, bone biology applications, skin
CC biology applications, cosmeceutical applications, endocrinology
CC applications, infectious disease applications, drug toxicities and drug
CC resistance applications, immunobiology, inflammation, and allergic
CC response applications, and biotechnology applications. The peptide
CC library can easily be monitored, both for its presence within cells and
CC its quantity. The expression of structurally biased libraries generate
CC elevated cellular concentration of peptides having a given structural
CC bias and thus increase the hit rate for targets that bind such
CC structures. This is the amino acid sequence of a scaffold protein used in
CC peptide libraries ot hold the library peptide in a conformationally
CC restricted form.
XX
SQ Sequence 83 AA;

Query Match 60.0%; Score 63; DB 7; Length 83;
Best Local Similarity 66.7%; Pred. No. 0.47;
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |: ||| ||| ||||| ||| |:
Db         57 AAKAAAEAAAKAAAEAAAKAAAAK 80
```

RESULT 34
ADE10645
ID ADE10645 standard; protein; 83 AA.
XX
AC ADE10645;
XX
DT 29-JAN-2004 (first entry)
XX
DE Structurally biased random peptide library related protein seqid 52.
XX
KW fusion nucleic acid library; scaffold protein; bioactive peptide;
KW phenotype change; cell morphology; cell growth; cell viability;
KW cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW loss of cell division; decreased cell growth; brca-1; brca-2;
KW tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW skin biology; cosmeceutical; endocrinology; infectious disease;
KW drug toxicity; drug resistance; inflammation; allergic response.
XX
OS Synthetic.
XX
PN US2003143562-A1.
XX
PD 31-JUL-2003.
XX
PF 20-JUN-2002; 2002US-00177725.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2003-829786/77.
XX
PT Novel library of fusion nucleic acids each of which has fused first and
PT second nucleic acids encoding scaffold protein and library peptide having
PT alpha helical biasing sequence, respectively, useful in screening
PT methods.
XX
PS Example 6; SEQ ID NO 52; 110pp; English.
XX
CC The invention describes a library (I) of fusion nucleic acids, where each
CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC library peptide sequence comprising an alpha helical biasing sequence;
CC where N1 is fused to N2. Disclosed is a method for screening bioactive
CC peptides conferring a change in specific phenotype such as cell
CC morphology, cell growth, cell viability, adhesion to substrates or other
CC cells, and cellular density; changes in the expression of one or more
CC RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC peptide identified by above mentioned method is used to generate more
CC candidate peptides and to identify target molecules, i.e., the molecules
CC with which the bioactive peptide interacts. The peptide(s) can be
CC combined with other pharmacologic activators to study the epistatic
CC relationships of signal transduction pathways in question. The disclosed

```
CC   method is also useful in cancer applications. Random libraries can be
CC   introduced into any tumour cell (primary or cultured), and peptides
CC   identified which by themselves induce apoptosis, cell death, loss of cell
CC   division or decreased cell growth. The method is also useful for
CC   screening of bioactive peptides which restore the constitutive function
CC   of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC   important in breast cancer such as the adenomatous polyposis coli gene
CC   (APC) and the Drosophila discs-large gene (DIg), which are components of
CC   cell-cell junctions. The methods are useful in cardiovascular
CC   applications, neurobiology applications, bone biology applications, skin
CC   biology applications, cosmeceutical applications, endocrinology
CC   applications, infectious disease applications, drug toxicities and drug
CC   resistance applications, immunobiology, inflammation, and allergic
CC   response applications, and biotechnology applications. The peptide
CC   library can easily be monitored, both for its presence within cells and
CC   its quantity. The expression of structurally biased libraries generate
CC   elevated cellular concentration of peptides having a given structural
CC   bias and thus increase the hit rate for targets that bind such
CC   structures. This is the amino acid sequence of a protein associated with
CC   fused nucleic acid and random peptide libraries of the invention.
XX
SQ   Sequence 83 AA;

  Query Match             60.0%;  Score 63;  DB 7;  Length 83;
  Best Local Similarity   66.7%;  Pred. No. 0.47;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         57 AAKAAAEAAAKAAAEAAAKAAAAK 80


RESULT 35
ADK15714
ID   ADK15714 standard; peptide; 83 AA.
XX
AC   ADK15714;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Library fusion protein-related scaffold protein #46.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2004-033956/03.
XX
PT   Library of fusion polypeptides in which each polypeptides comprises
PT   scaffold protein and library peptide having alpha helical biasing
PT   sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS   Disclosure; SEQ ID NO 102; 110pp; English.
XX
CC   The invention comprises a library of fusion nucleic acids, where each
CC   encoded protein contains a scaffold protein (e.g. a green fluorescent
CC   protein - GFP) and a library peptide sequence comprising an alpha helical
CC   biasing sequence, or a scaffold protein, a library peptide and a
CC   nucleating sequence. The library of the invention is useful for screening
CC   bioactive peptides conferring a particular phenotype. The present amino
CC   acid sequence represents a scaffold protein.
XX
SQ   Sequence 83 AA;

  Query Match             60.0%;  Score 63;  DB 8;  Length 83;
  Best Local Similarity   66.7%;  Pred. No. 0.47;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         57 AAKAAAEAAAKAAAEAAAKAAAAK 80


RESULT 36
ADK15664
ID   ADK15664 standard; peptide; 83 AA.
XX
AC   ADK15664;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Nucleating sequence-containing library fusion protein #46.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2004-033956/03.
XX
PT   Library of fusion polypeptides in which each polypeptides comprises
PT   scaffold protein and library peptide having alpha helical biasing
PT   sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS   Example 6; SEQ ID NO 52; 110pp; English.
XX
CC   The invention comprises a library of fusion nucleic acids, where each
CC   encoded protein contains a scaffold protein (e.g. a green fluorescent
CC   protein - GFP) and a library peptide sequence comprising an alpha helical
CC   biasing sequence, or a scaffold protein, a library peptide and a
CC   nucleating sequence. The library of the invention is useful for screening
CC   bioactive peptides conferring a particular phenotype. The present amino
CC   acid sequence represents a library protein containing a nucleating
CC   sequence.
XX
SQ   Sequence 83 AA;

  Query Match             60.0%;  Score 63;  DB 8;  Length 83;
  Best Local Similarity   66.7%;  Pred. No. 0.47;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;
```

```
Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         57 AAKAAAEAAAKAAAEAAAKAAAAK 80


RESULT 37
ADE10642
ID   ADE10642 standard; protein; 88 AA.
XX
AC   ADE10642;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Structurally biased random peptide library related protein seqid 49.
XX
KW   fusion nucleic acid library; scaffold protein; bioactive peptide;
KW   phenotype change; cell morphology; cell growth; cell viability;
KW   cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW   loss of cell division; decreased cell growth; brca-1; brca-2;
KW   tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW   Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW   skin biology; cosmeceutical; endocrinology; infectious disease;
KW   drug toxicity; drug resistance; inflammation; allergic response.
XX
OS   Synthetic.
XX
PN   US2003143562-A1.
XX
PD   31-JUL-2003.
XX
PF   20-JUN-2002; 2002US-00177725.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2003-829786/77.
XX
PT   Novel library of fusion nucleic acids each of which has fused first and
PT   second nucleic acids encoding scaffold protein and library peptide having
PT   alpha helical biasing sequence, respectively, useful in screening
PT   methods.
XX
PS   Example 6; SEQ ID NO 49; 110pp; English.
XX
CC   The invention describes a library (I) of fusion nucleic acids, where each
CC   fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC   scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC   library peptide sequence comprising an alpha helical biasing sequence;
CC   where N1 is fused to N2. Disclosed is a method for screening bioactive
CC   peptides conferring a change in specific phenotype such as cell
CC   morphology, cell growth, cell viability, adhesion to substrates or other
CC   cells, and cellular density; changes in the expression of one or more
CC   RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC   in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC   lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC   peptide identified by above mentioned method is used to generate more
CC   candidate peptides and to identify target molecules, i.e., the molecules
CC   with which the bioactive peptide interacts. The peptide(s) can be
CC   combined with other pharmacologic activators to study the epistatic
CC   relationships of signal transduction pathways in question. The disclosed
CC   method is also useful in cancer applications. Random libraries can be
CC   introduced into any tumour cell (primary or cultured), and peptides
CC   identified which by themselves induce apoptosis, cell death, loss of cell
CC   division or decreased cell growth. The method is also useful for
CC   screening of bioactive peptides which restore the constitutive function
CC   of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC   important in breast cancer such as the adenomatous polyposis coli gene
CC   (APC) and the Drosophila discs-large gene (DIg), which are components of
CC   cell-cell junctions. The methods are useful in cardiovascular
CC   applications, neurobiology applications, bone biology applications, skin
CC   biology applications, cosmeceutical applications, endocrinology
CC   applications, infectious disease applications, drug toxicities and drug
CC   resistance applications, immunobiology, inflammation, and allergic
CC   response applications, and biotechnology applications. The peptide
CC   library can easily be monitored, both for its presence within cells and
CC   its quantity. The expression of structurally biased libraries generate
CC   elevated cellular concentration of peptides having a given structural
CC   bias and thus increase the hit rate for targets that bind such
CC   structures. This is the amino acid sequence of a protein associated with
CC   fused nucleic acid and random peptide libraries of the invention.
XX
SQ   Sequence 88 AA;

  Query Match             60.0%;  Score 63;  DB 7;  Length 88;
  Best Local Similarity   66.7%;  Pred. No. 0.5;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         63 AAKAAAEAAAKAAAEAAAKAAAAK 86


RESULT 38
ADE10692
ID   ADE10692 standard; protein; 88 AA.
XX
AC   ADE10692;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Structurally biased random peptide library scaffold protein seqid 99.
XX
KW   fusion nucleic acid library; scaffold protein; bioactive peptide;
KW   phenotype change; cell morphology; cell growth; cell viability;
KW   cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW   loss of cell division; decreased cell growth; brca-1; brca-2;
KW   tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW   Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW   skin biology; cosmeceutical; endocrinology; infectious disease;
KW   drug toxicity; drug resistance; inflammation; allergic response;
KW   scaffold protein.
XX
OS   Synthetic.
XX
PN   US2003143562-A1.
XX
PD   31-JUL-2003.
XX
PF   20-JUN-2002; 2002US-00177725.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2003-829786/77.
XX
PT   Novel library of fusion nucleic acids each of which has fused first and
PT   second nucleic acids encoding scaffold protein and library peptide having
PT   alpha helical biasing sequence, respectively, useful in screening
PT   methods.
XX
PS   Disclosure; SEQ ID NO 99; 110pp; English.
XX
CC   The invention describes a library (I) of fusion nucleic acids, where each
CC   fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC   scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC   library peptide sequence comprising an alpha helical biasing sequence;
CC   where N1 is fused to N2. Disclosed is a method for screening bioactive
CC   peptides conferring a change in specific phenotype such as cell
```

```
CC   morphology, cell growth, cell viability, adhesion to substrates or other
CC   cells, and cellular density; changes in the expression of one or more
CC   RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC   in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC   lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC   peptide identified by above mentioned method is used to generate more
CC   candidate peptides and to identify target molecules, i.e., the molecules
CC   with which the bioactive peptide interacts. The peptide(s) can be
CC   combined with other pharmacologic activators to study the epistatic
CC   relationships of signal transduction pathways in question. The disclosed
CC   method is also useful in cancer applications. Random libraries can be
CC   introduced into any tumour cell (primary or cultured), and peptides
CC   identified which by themselves induce apoptosis, cell death, loss of cell
CC   division or decreased cell growth. The method is also useful for
CC   screening of bioactive peptides which restore the constitutive function
CC   of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC   important in breast cancer such as the adenomatous polyposis coli gene
CC   (APC) and the Drosophila discs-large gene (DIg), which are components of
CC   cell-cell junctions. The methods are useful in cardiovascular
CC   applications, neurobiology applications, bone biology applications, skin
CC   biology applications, cosmeceutical applications, endocrinology
CC   applications, infectious disease applications, drug toxicities and drug
CC   resistance applications, immunobiology, inflammation, and allergic
CC   response applications, and biotechnology applications. The peptide
CC   library can easily be monitored, both for its presence within cells and
CC   its quantity. The expression of structurally biased libraries generate
CC   elevated cellular concentration of peptides having a given structural
CC   bias and thus increase the hit rate for targets that bind such
CC   structures. This is the amino acid sequence of a scaffold protein used in
CC   peptide libraries ot hold the library peptide in a conformationally
CC   restricted form.
XX
SQ   Sequence 88 AA;

  Query Match             60.0%;  Score 63;  DB 7;  Length 88;
  Best Local Similarity   66.7%;  Pred. No. 0.5;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         63 AAKAAAEAAAKAAAEAAAKAAAAK 86


RESULT 39
ADK15711
ID   ADK15711 standard; peptide; 88 AA.
XX
AC   ADK15711;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Library fusion protein-related scaffold protein #43.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2004-033956/03.
XX
PT   Library of fusion polypeptides in which each polypeptides comprises
PT   scaffold protein and library peptide having alpha helical biasing
PT   sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS   Disclosure; SEQ ID NO 99; 110pp; English.
XX
CC   The invention comprises a library of fusion nucleic acids, where each
CC   encoded protein contains a scaffold protein (e.g. a green fluorescent
CC   protein - GFP) and a library peptide sequence comprising an alpha helical
CC   biasing sequence, or a scaffold protein, a library peptide and a
CC   nucleating sequence. The library of the invention is useful for screening
CC   bioactive peptides conferring a particular phenotype. The present amino
CC   acid sequence represents a scaffold protein.
XX
SQ   Sequence 88 AA;

  Query Match             60.0%;  Score 63;  DB 8;  Length 88;
  Best Local Similarity   66.7%;  Pred. No. 0.5;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         63 AAKAAAEAAAKAAAEAAAKAAAAK 86


RESULT 40
ADK15661
ID   ADK15661 standard; peptide; 88 AA.
XX
AC   ADK15661;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Nucleating sequence-containing library fusion protein #43.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2004-033956/03.
XX
PT   Library of fusion polypeptides in which each polypeptides comprises
PT   scaffold protein and library peptide having alpha helical biasing
PT   sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS   Example 6; SEQ ID NO 49; 110pp; English.
XX
CC   The invention comprises a library of fusion nucleic acids, where each
CC   encoded protein contains a scaffold protein (e.g. a green fluorescent
CC   protein - GFP) and a library peptide sequence comprising an alpha helical
CC   biasing sequence, or a scaffold protein, a library peptide and a
```

```
CC   nucleating sequence. The library of the invention is useful for screening
CC   bioactive peptides conferring a particular phenotype. The present amino
CC   acid sequence represents a library protein containing a nucleating
CC   sequence.
XX
SQ   Sequence 88 AA;

  Query Match             60.0%;  Score 63;  DB 8;  Length 88;
  Best Local Similarity   66.7%;  Pred. No. 0.5;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         63 AAKAAAEAAAKAAAEAAAKAAAAK 86


RESULT 41
ADE10694
ID   ADE10694 standard; protein; 91 AA.
XX
AC   ADE10694;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Structurally biased random peptide library scaffold protein seqid 101.
XX
KW   fusion nucleic acid library; scaffold protein; bioactive peptide;
KW   phenotype change; cell morphology; cell growth; cell viability;
KW   cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW   loss of cell division; decreased cell growth; brca-1; brca-2;
KW   tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW   Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW   skin biology; cosmeceutical; endocrinology; infectious disease;
KW   drug toxicity; drug resistance; inflammation; allergic response;
KW   scaffold protein.
XX
OS   Synthetic.
XX
PN   US2003143562-A1.
XX
PD   31-JUL-2003.
XX
PF   20-JUN-2002; 2002US-00177725.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2003-829786/77.
XX
PT   Novel library of fusion nucleic acids each of which has fused first and
PT   second nucleic acids encoding scaffold protein and library peptide having
PT   alpha helical biasing sequence, respectively, useful in screening
PT   methods.
XX
PS   Disclosure; SEQ ID NO 101; 110pp; English.
XX
CC   The invention describes a library (I) of fusion nucleic acids, where each
CC   fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC   scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC   library peptide sequence comprising an alpha helical biasing sequence;
CC   where N1 is fused to N2. Disclosed is a method for screening bioactive
CC   peptides conferring a change in specific phenotype such as cell
CC   morphology, cell growth, cell viability, adhesion to substrates or other
CC   cells, and cellular density; changes in the expression of one or more
CC   RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC   in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC   lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC   peptide identified by above mentioned method is used to generate more
CC   candidate peptides and to identify target molecules, i.e., the molecules
CC   with which the bioactive peptide interacts. The peptide(s) can be
CC   combined with other pharmacologic activators to study the epistatic
CC   relationships of signal transduction pathways in question. The disclosed
CC   method is also useful in cancer applications. Random libraries can be
CC   introduced into any tumour cell (primary or cultured), and peptides
CC   identified which by themselves induce apoptosis, cell death, loss of cell
CC   division or decreased cell growth. The method is also useful for
CC   screening of bioactive peptides which restore the constitutive function
CC   of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC   important in breast cancer such as the adenomatous polyposis coli gene
CC   (APC) and the Drosophila discs-large gene (DIg), which are components of
CC   cell-cell junctions. The methods are useful in cardiovascular
CC   applications, neurobiology applications, bone biology applications, skin
CC   biology applications, cosmeceutical applications, endocrinology
CC   applications, infectious disease applications, drug toxicities and drug
CC   resistance applications, immunobiology, inflammation, and allergic
CC   response applications, and biotechnology applications. The peptide
CC   library can easily be monitored, both for its presence within cells and
CC   its quantity. The expression of structurally biased libraries generate
CC   elevated cellular concentration of peptides having a given structural
CC   bias and thus increase the hit rate for targets that bind such
CC   structures. This is the amino acid sequence of a scaffold protein used in
CC   peptide libraries ot hold the library peptide in a conformationally
CC   restricted form.
XX
SQ   Sequence 91 AA;

  Query Match             60.0%;  Score 63;  DB 7;  Length 91;
  Best Local Similarity   66.7%;  Pred. No. 0.52;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         65 AAKAAAEAAAKAAAEAAAKAAAAK 88


RESULT 42
ADE10644
ID   ADE10644 standard; protein; 91 AA.
XX
AC   ADE10644;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Structurally biased random peptide library related protein seqid 51.
XX
KW   fusion nucleic acid library; scaffold protein; bioactive peptide;
KW   phenotype change; cell morphology; cell growth; cell viability;
KW   cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW   loss of cell division; decreased cell growth; brca-1; brca-2;
KW   tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW   Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW   skin biology; cosmeceutical; endocrinology; infectious disease;
KW   drug toxicity; drug resistance; inflammation; allergic response.
XX
OS   Synthetic.
XX
PN   US2003143562-A1.
XX
PD   31-JUL-2003.
XX
PF   20-JUN-2002; 2002US-00177725.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2003-829786/77.
XX
PT   Novel library of fusion nucleic acids each of which has fused first and
```

```
PT   second nucleic acids encoding scaffold protein and library peptide having
PT   alpha helical biasing sequence, respectively, useful in screening
PT   methods.
XX
PS   Example 6; SEQ ID NO 51; 110pp; English.
XX
CC   The invention describes a library (I) of fusion nucleic acids, where each
CC   fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC   scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC   library peptide sequence comprising an alpha helical biasing sequence;
CC   where N1 is fused to N2. Disclosed is a method for screening bioactive
CC   peptides conferring a change in specific phenotype such as cell
CC   morphology, cell growth, cell viability, adhesion to substrates or other
CC   cells, and cellular density; changes in the expression of one or more
CC   RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC   in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC   lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC   peptide identified by above mentioned method is used to generate more
CC   candidate peptides and to identify target molecules, i.e., the molecules
CC   with which the bioactive peptide interacts. The peptide(s) can be
CC   combined with other pharmacologic activators to study the epistatic
CC   relationships of signal transduction pathways in question. The disclosed
CC   method is also useful in cancer applications. Random libraries can be
CC   introduced into any tumour cell (primary or cultured), and peptides
CC   identified which by themselves induce apoptosis, cell death, loss of cell
CC   division or decreased cell growth. The method is also useful for
CC   screening of bioactive peptides which restore the constitutive function
CC   of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC   important in breast cancer such as the adenomatous polyposis coli gene
CC   (APC) and the Drosophila discs-large gene (DIg), which are components of
CC   cell-cell junctions. The methods are useful in cardiovascular
CC   applications, neurobiology applications, bone biology applications, skin
CC   biology applications, cosmeceutical applications, endocrinology
CC   applications, infectious disease applications, drug toxicities and drug
CC   resistance applications, immunobiology, inflammation, and allergic
CC   response applications, and biotechnology applications. The peptide
CC   library can easily be monitored, both for its presence within cells and
CC   its quantity. The expression of structurally biased libraries generate
CC   elevated cellular concentration of peptides having a given structural
CC   bias and thus increase the hit rate for targets that bind such
CC   structures. This is the amino acid sequence of a protein associated with
CC   fused nucleic acid and random peptide libraries of the invention.
XX
SQ   Sequence 91 AA;

  Query Match             60.0%;  Score 63;  DB 7;  Length 91;
  Best Local Similarity   66.7%;  Pred. No. 0.52;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         65 AAKAAAEAAAKAAAEAAAKAAAAK 88


RESULT 43
ADK15663
ID   ADK15663 standard; peptide; 91 AA.
XX
AC   ADK15663;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Nucleating sequence-containing library fusion protein #45.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
DR   WPI; 2004-033956/03.
XX
PT   Library of fusion polypeptides in which each polypeptides comprises
PT   scaffold protein and library peptide having alpha helical biasing
PT   sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS   Example 6; SEQ ID NO 51; 110pp; English.
XX
CC   The invention comprises a library of fusion nucleic acids, where each
CC   encoded protein contains a scaffold protein (e.g. a green fluorescent
CC   protein - GFP) and a library peptide sequence comprising an alpha helical
CC   biasing sequence, or a scaffold protein, a library peptide and a
CC   nucleating sequence. The library of the invention is useful for screening
CC   bioactive peptides conferring a particular phenotype. The present amino
CC   acid sequence represents a library protein containing a nucleating
CC   sequence.
XX
SQ   Sequence 91 AA;

  Query Match             60.0%;  Score 63;  DB 8;  Length 91;
  Best Local Similarity   66.7%;  Pred. No. 0.52;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         65 AAKAAAEAAAKAAAEAAAKAAAAK 88


RESULT 44
ADK15713
ID   ADK15713 standard; peptide; 91 AA.
XX
AC   ADK15713;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Library fusion protein-related scaffold protein #45.
XX
KW   fusion nucleic acid library; fusion protein library; scaffold protein;
KW   green fluorescent protein; GFP; alpha helical biasing sequence;
KW   nucleating sequence; screening.
XX
OS   Synthetic.
XX
PN   US2003224412-A1.
XX
PD   04-DEC-2003.
XX
PF   18-MAR-2003; 2003US-00393449.
XX
PR   08-OCT-1998;   98US-00169015.
PR   08-OCT-1999;   99US-00415765.
PR   20-JUN-2002; 2002US-00177725.
XX
PA   (ANDE/) ANDERSON D.
PA   (PEEL/) PEELLE B R.
PA   (BOGE/) BOGENBERGER J M.
XX
PI   Anderson D,  Peelle BR,  Bogenberger JM;
XX
```

DR WPI; 2004-033956/03.
XX
PT Library of fusion polypeptides in which each polypeptides comprises
PT scaffold protein and library peptide having alpha helical biasing
PT sequence, or scaffold protein, library peptide and nucleating sequence.
XX
PS Disclosure; SEQ ID NO 101; 110pp; English.
XX
CC The invention comprises a library of fusion nucleic acids, where each
CC encoded protein contains a scaffold protein (e.g. a green fluorescent
CC protein - GFP) and a library peptide sequence comprising an alpha helical
CC biasing sequence, or a scaffold protein, a library peptide and a
CC nucleating sequence. The library of the invention is useful for screening
CC bioactive peptides conferring a particular phenotype. The present amino
CC acid sequence represents a scaffold protein.
XX
SQ Sequence 91 AA;

Query Match 60.0%; Score 63; DB 8; Length 91;
Best Local Similarity 66.7%; Pred. No. 0.52;
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         65 AAKAAAEAAAKAAAEAAAKAAAAK 88
```

RESULT 45
ADE10690
ID ADE10690 standard; protein; 104 AA.
XX
AC ADE10690;
XX
DT 29-JAN-2004 (first entry)
XX
DE Structurally biased random peptide library scaffold protein seqid 97.
XX
KW fusion nucleic acid library; scaffold protein; bioactive peptide;
KW phenotype change; cell morphology; cell growth; cell viability;
KW cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;
KW loss of cell division; decreased cell growth; brca-1; brca-2;
KW tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;
KW Drosophila discs-large; DIg; cardiovascular; neurobiology; bone biology;
KW skin biology; cosmeceutical; endocrinology; infectious disease;
KW drug toxicity; drug resistance; inflammation; allergic response;
KW scaffold protein.
XX
OS Synthetic.
XX
PN US2003143562-A1.
XX
PD 31-JUL-2003.
XX
PF 20-JUN-2002; 2002US-00177725.
XX
PR 08-OCT-1998; 98US-00169015.
PR 08-OCT-1999; 99US-00415765.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Anderson D, Peelle BR, Bogenberger JM;
XX
DR WPI; 2003-829786/77.
XX
PT Novel library of fusion nucleic acids each of which has fused first and
PT second nucleic acids encoding scaffold protein and library peptide having
PT alpha helical biasing sequence, respectively, useful in screening
PT methods.
XX
PS Disclosure; SEQ ID NO 97; 110pp; English.
XX
CC The invention describes a library (I) of fusion nucleic acids, where each
CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a
CC scaffold protein sequence; and a second nucleic acid (N2), encoding a
CC library peptide sequence comprising an alpha helical biasing sequence;
CC where N1 is fused to N2. Disclosed is a method for screening bioactive
CC peptides conferring a change in specific phenotype such as cell
CC morphology, cell growth, cell viability, adhesion to substrates or other
CC cells, and cellular density; changes in the expression of one or more
CC RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes
CC in the equilibrium state (i.e., half-life) or one or more RNAs, protein,
CC lipids, hormones, cytokines, or other molecules; etc. The bioactive
CC peptide identified by above mentioned method is used to generate more
CC candidate peptides and to identify target molecules, i.e., the molecules
CC with which the bioactive peptide interacts. The peptide(s) can be
CC combined with other pharmacologic activators to study the epistatic
CC relationships of signal transduction pathways in question. The disclosed
CC method is also useful in cancer applications. Random libraries can be
CC introduced into any tumour cell (primary or cultured), and peptides
CC identified which by themselves induce apoptosis, cell death, loss of cell
CC division or decreased cell growth. The method is also useful for
CC screening of bioactive peptides which restore the constitutive function
CC of the brca-1 or brca-2 genes, and other tumour suppressor genes
CC important in breast cancer such as the adenomatous polyposis coli gene
CC (APC) and the Drosophila discs-large gene (DIg), which are components of
CC cell-cell junctions. The methods are useful in cardiovascular
CC applications, neurobiology applications, bone biology applications, skin
CC biology applications, cosmeceutical applications, endocrinology
CC applications, infectious disease applications, drug toxicities and drug
CC resistance applications, immunobiology, inflammation, and allergic
CC response applications, and biotechnology applications. The peptide
CC library can easily be monitored, both for its presence within cells and
CC its quantity. The expression of structurally biased libraries generate
CC elevated cellular concentration of peptides having a given structural
CC bias and thus increase the hit rate for targets that bind such
CC structures. This is the amino acid sequence of a scaffold protein used in
CC peptide libraries ot hold the library peptide in a conformationally
CC restricted form.
XX
SQ Sequence 104 AA;

Query Match 60.0%; Score 63; DB 7; Length 104;
Best Local Similarity 66.7%; Pred. No. 0.61;
Matches 16; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         79 AAKAAAEAAAKAAAEAAAKAAAAK 102
```

Search completed: July 11, 2005, 09:43:30
Job time : 163 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2005, 09:37:04 ; Search time 42 Seconds
(without alignments)
44.434 Million cell updates/sec

Title: SEQ1
Perfect score: 105
Sequence: 1 axaeaaekaakyaaeaaekaakaxa 25

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	68	64.8	69	4	US-09-902-540-14824	Sequence 14824, A
2	61.5	58.6	56	4	US-09-405-743A-3	Sequence 3, Appli
3	61.5	58.6	56	4	US-09-816-989A-3	Sequence 3, Appli
4	61.5	58.6	86	4	US-09-405-743A-6	Sequence 6, Appli
5	61.5	58.6	86	4	US-09-816-989A-6	Sequence 6, Appli
6	61	58.1	33	1	US-08-303-025-16	Sequence 16, Appl
7	61	58.1	33	2	US-08-436-703B-4	Sequence 4, Appli
8	61	58.1	469	4	US-09-489-039A-13565	Sequence 13565, A
9	60.5	57.6	117	4	US-09-340-736E-9	Sequence 9, Appli
10	60.5	57.6	117	4	US-09-964-662-9	Sequence 9, Appli
11	60.5	57.6	118	4	US-09-340-736E-10	Sequence 10, Appl
12	60.5	57.6	118	4	US-09-964-662-10	Sequence 10, Appl
13	60.5	57.6	199	4	US-09-340-736E-11	Sequence 11, Appl
14	60.5	57.6	199	4	US-09-964-662-11	Sequence 11, Appl
15	60.5	57.6	200	4	US-09-340-736E-2	Sequence 2, Appli
16	60.5	57.6	200	4	US-09-964-662-2	Sequence 2, Appli
17	60.5	57.6	201	2	US-08-911-364-2	Sequence 2, Appli
18	60.5	57.6	730	4	US-09-961-403-8	Sequence 8, Appli
19	60.5	57.6	731	2	US-08-911-364-1	Sequence 1, Appli
20	60.5	57.6	731	4	US-09-340-736E-1	Sequence 1, Appli
21	60.5	57.6	731	4	US-09-964-662-1	Sequence 1, Appli
22	60.5	57.6	733	3	US-08-464-700-2	Sequence 2, Appli
23	59.5	56.7	67	4	US-09-869-875-7	Sequence 7, Appli
24	59	56.2	45	4	US-09-405-743A-2	Sequence 2, Appli
25	59	56.2	45	4	US-09-816-989A-2	Sequence 2, Appli
26	59	56.2	109	4	US-09-405-743A-7	Sequence 7, Appli
27	59	56.2	109	4	US-09-816-989A-7	Sequence 7, Appli
28	58	55.2	77	4	US-09-405-743A-5	Sequence 5, Appli
29	58	55.2	77	4	US-09-816-989A-5	Sequence 5, Appli
30	58	55.2	92	4	US-09-344-529-2	Sequence 2, Appli
31	57.5	54.8	325	4	US-09-902-540-13678	Sequence 13678, A
32	57	54.3	28	1	US-08-303-025-12	Sequence 12, Appl
33	57	54.3	28	2	US-08-436-703B-1	Sequence 1, Appli
34	57	54.3	29	1	US-08-152-488-10	Sequence 10, Appl
35	57	54.3	29	1	US-08-152-488-11	Sequence 11, Appl
36	57	54.3	29	1	US-08-152-488-12	Sequence 12, Appl
37	57	54.3	29	1	US-08-303-025-10	Sequence 10, Appl
38	57	54.3	29	1	US-08-303-025-11	Sequence 11, Appl
39	57	54.3	29	1	US-08-303-025-13	Sequence 13, Appl
40	57	54.3	29	1	US-08-303-025-14	Sequence 14, Appl
41	57	54.3	29	1	US-08-677-304-10	Sequence 10, Appl
42	57	54.3	29	1	US-08-677-304-11	Sequence 11, Appl
43	57	54.3	29	1	US-08-677-304-12	Sequence 12, Appl
44	57	54.3	29	2	US-08-436-703B-3	Sequence 3, Appli
45	57	54.3	29	2	US-08-436-703B-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-902-540-14824
; Sequence 14824, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14824
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14824

Query Match 64.8%; Score 68; DB 4; Length 69;
Best Local Similarity 64.0%; Pred. No. 0.018;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:  ||| ||| ||||| :||:|:|
Db         12 AAKRAAEAAAKRAAEAARRAAEAAA 36
```

RESULT 2
US-09-405-743A-3
; Sequence 3, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co., Ltd.
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
; FILE REFERENCE: 60807-A
; CURRENT APPLICATION NUMBER: US/09/405,743A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE

```
US-09-405-743A-3

  Query Match             58.6%;  Score 61.5;  DB 4;  Length 56;
  Best Local Similarity   64.0%;  Pred. No. 0.1;
  Matches   16;  Conservative    4;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          2 XAEAAEKA-AKYAAEAAEKAAKAXA 25
              :||| :|| ||| |||| ||| :|
Db         29 AAEAKKKAEAKYKAEAAKAAAKEAA 53


RESULT 3
US-09-816-989A-3
; Sequence 3, Application US/09816989A
; Patent No. 6800287
; GENERAL INFORMATION:
;  APPLICANT: Gad, Alexander
;  APPLICANT:  Lis, Doris
;  TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
;  TITLE OF INVENTION:  AND FOR THERAPEUTIC USE
;  FILE REFERENCE: 2609/60807-A-PCT-US
;  CURRENT APPLICATION NUMBER: US/09/816,989A
;  CURRENT FILING DATE:  2001-03-23
;  PRIOR APPLICATION NUMBER: 60/101,693
;  PRIOR FILING DATE: 1998-09-25
;  PRIOR APPLICATION NUMBER: PCT/US99/22402
;  PRIOR FILING DATE: 1999-09-24
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 56
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-3

  Query Match             58.6%;  Score 61.5;  DB 4;  Length 56;
  Best Local Similarity   64.0%;  Pred. No. 0.1;
  Matches   16;  Conservative    4;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          2 XAEAAEKA-AKYAAEAAEKAAKAXA 25
              :||| :|| ||| |||| ||| :|
Db         29 AAEAKKKAEAKYKAEAAKAAAKEAA 53


RESULT 4
US-09-405-743A-6
; Sequence 6, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
;  APPLICANT:  Yeda Research and Development Co., Ltd.
;  TITLE OF INVENTION:  GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
;  FILE REFERENCE:  60807-A
;  CURRENT APPLICATION NUMBER: US/09/405,743A
;  CURRENT FILING DATE:  1999-09-24
;  NUMBER OF SEQ ID NOS:  7
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 86
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:   SYNTHETIC
;   OTHER INFORMATION: PEPTIDE
US-09-405-743A-6

  Query Match             58.6%;  Score 61.5;  DB 4;  Length 86;
  Best Local Similarity   60.7%;  Pred. No. 0.17;
  Matches   17;  Conservative    3;  Mismatches    5;  Indels    3;  Gaps    1;

Qy          1 AXAEAAEKA---AKYAAEAAEKAAKAXA 25
              |:||  | |   ||| ||||:|| ||:|
Db         47 AKAEKKEYAAAEAKYKAEAAKKAYKAEA 74


RESULT 5
US-09-816-989A-6
; Sequence 6, Application US/09816989A
; Patent No. 6800287
; GENERAL INFORMATION:
;  APPLICANT: Gad, Alexander
;  APPLICANT:  Lis, Doris
;  TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARKE
;  TITLE OF INVENTION:  AND FOR THERAPEUTIC USE
;  FILE REFERENCE: 2609/60807-A-PCT-US
;  CURRENT APPLICATION NUMBER: US/09/816,989A
;  CURRENT FILING DATE:  2001-03-23
;  PRIOR APPLICATION NUMBER: 60/101,693
;  PRIOR FILING DATE: 1998-09-25
;  PRIOR APPLICATION NUMBER: PCT/US99/22402
;  PRIOR FILING DATE: 1999-09-24
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 86
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-6

  Query Match             58.6%;  Score 61.5;  DB 4;  Length 86;
  Best Local Similarity   60.7%;  Pred. No. 0.17;
  Matches   17;  Conservative    3;  Mismatches    5;  Indels    3;  Gaps    1;

Qy          1 AXAEAAEKA---AKYAAEAAEKAAKAXA 25
              |:||  | |   ||| ||||:|| ||:|
Db         47 AKAEKKEYAAAEAKYKAEAAKKAYKAEA 74


RESULT 6
US-08-303-025-16
; Sequence 16, Application US/08303025
; Patent No. 5614494
;  GENERAL INFORMATION:
;    APPLICANT:  Wakefield, Thomas W.
;    APPLICANT:  Andrews, Philip C.
;    APPLICANT:  Stanley, James C.
;    TITLE OF INVENTION:  NOVEL PEPTIDES FOR HEPARIN AND
;    TITLE OF INVENTION:  LOW MOLECULAR WEIGHT HEPARIN
;    TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benita J, Rohm, Esq.
;      STREET:  150 West Jefferson, Suite 2500
;      CITY:  Detroit
;      STATE:  Michigan
;      COUNTRY:  United States of America
;      ZIP:  48226-4415
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy diskette 3.5"  1.44Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS v.6.22
;      SOFTWARE:  WordPerfect  6.1; ASCII (DOS)Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/303,025
;      FILING DATE:  08-SEPT-1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06829
;      FILING DATE:  14-AUG-1992
;      APPLICATION NUMBER:  US 08/152,488
;      FILING DATE:  12-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rohm, Benita J.
```

```
;      REFERENCE/DOCKET NUMBER:  7WH-060548-00231
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  313-496-7622
;      TELEFAX:  313-496-8454
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  33 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  N/A
;      TOPOLOGY:  N/A
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  N/A
;    PUBLICATION INFORMATION:
;      AUTHORS:  N/A
;      TITLE:  N/A
;      DOCUMENT NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
US-08-303-025-16

  Query Match             58.1%;  Score 61;  DB 1;  Length 33;
  Best Local Similarity   66.7%;  Pred. No. 0.067;
  Matches   14;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 EAAEKAAKYAAEAAEKAAKAX 24
              |||:|||| | :||:|| ||:
Db          1 EAAKKAAKKAKKAAKKAKKAA 21


RESULT 7
US-08-436-703B-4
; Sequence 4, Application US/08436703B
; Patent No. 5919761
;  GENERAL INFORMATION:
;    APPLICANT:  Wakefield, Thomas W.
;    APPLICANT:  Andrews, Philip C.
;    APPLICANT:  Stanley, James C.
;    TITLE OF INVENTION:  NOVEL PEPTIDES FOR
;    TITLE OF INVENTION:  HEPARIN AND LOW MOLECULAR
;    TITLE OF INVENTION:  WEIGHT HEPARIN
;    TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benita J, Rohm, Esq.
;      STREET:  6601 Woodward Avenue
;      STREET:  Suite 1525
;      CITY:  Detroit
;      STATE:  Michigan
;      COUNTRY:  United States of America
;      ZIP:  48226
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk 1.44Mb, 3.5"
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 6;
;      SOFTWARE:  ASCII (DOS)Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/436,703B
;      FILING DATE:  08-MAY-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  N/A
;      FILING DATE:  N/A
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rohm, Benita J.
;      REGISTRATION NUMBER:  28,664
;      REFERENCE/DOCKET NUMBER:  7WK-060548-00233
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  313-965-1976
;      TELEFAX:  313-965-1951
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  33 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  N/A
;      TOPOLOGY:  N/A
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  N/A
;    PUBLICATION INFORMATION:
;      AUTHORS:  N/A
;      TITLE:  N/A
US-08-436-703B-4

  Query Match             58.1%;  Score 61;  DB 2;  Length 33;
  Best Local Similarity   66.7%;  Pred. No. 0.067;
  Matches   14;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 EAAEKAAKYAAEAAEKAAKAX 24
              |||:|||| | :||:|| ||:
Db          1 EAAKKAAKKAKKAAKKAKKAA 21


RESULT 8
US-09-489-039A-13565
; Sequence 13565, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 13565
;   LENGTH: 469
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13565

  Query Match             58.1%;  Score 61;  DB 4;  Length 469;
  Best Local Similarity   62.5%;  Pred. No. 1.3;
  Matches   15;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||||||     ||:|||||:|
Db        301 KAAAAEKAAAAEKAAADKAAKAAA 324


RESULT 9
US-09-340-736E-9
; Sequence 9, Application US/09340736E
; Patent No. 6489446
; GENERAL INFORMATION:
;  APPLICANT: ROTHSTEIN, ASER
;  APPLICANT:  KEELEY, FRED
;  APPLICANT:  ROTHSTEIN, STEVEN
;  TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN ELASTIN
;  TITLE OF INVENTION:  AND OTHER FIBROUS PROTEINS
;  FILE REFERENCE: 041082/0110
;  CURRENT APPLICATION NUMBER: US/09/340,736E
;  CURRENT FILING DATE:  1999-06-29
;  PRIOR APPLICATION NUMBER: 08/911,364
;  PRIOR FILING DATE: 1997-08-07
;  PRIOR APPLICATION NUMBER: 60/023,552
;  PRIOR FILING DATE: 1996-08-07
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 117
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
```

```
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: MFU-3 polypeptide
US-09-340-736E-9

  Query Match             57.6%;  Score 60.5;  DB 4;  Length 117;
  Best Local Similarity   58.6%;  Pred. No. 0.32;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db         37 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 65


RESULT 10
US-09-964-662-9
; Sequence 9, Application US/09964662
; Patent No. 6765086
; GENERAL INFORMATION:
;  APPLICANT: PROTEIN SPECIALTIES LTD.
;  APPLICANT:  HSC RESEARCH AND DEVELOPMENT LIMITED PARTNERSHIP
;  TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN ELASTIN AND
;  TITLE OF INVENTION:  OTHER FIBROUS PROTEINS
;  FILE REFERENCE: 041082/0112
;  CURRENT APPLICATION NUMBER: US/09/964,662
;  CURRENT FILING DATE:  2003-05-08
;  PRIOR APPLICATION NUMBER: 09/340,736
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 117
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-964-662-9

  Query Match             57.6%;  Score 60.5;  DB 4;  Length 117;
  Best Local Similarity   58.6%;  Pred. No. 0.32;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db         37 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 65


RESULT 11
US-09-340-736E-10
; Sequence 10, Application US/09340736E
; Patent No. 6489446
; GENERAL INFORMATION:
;  APPLICANT: ROTHSTEIN, ASER
;  APPLICANT:  KEELEY, FRED
;  APPLICANT:  ROTHSTEIN, STEVEN
;  TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN ELASTIN
;  TITLE OF INVENTION:  AND OTHER FIBROUS PROTEINS
;  FILE REFERENCE: 041082/0110
;  CURRENT APPLICATION NUMBER: US/09/340,736E
;  CURRENT FILING DATE:  1999-06-29
;  PRIOR APPLICATION NUMBER: 08/911,364
;  PRIOR FILING DATE: 1997-08-07
;  PRIOR APPLICATION NUMBER: 60/023,552
;  PRIOR FILING DATE: 1996-08-07
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 118
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: MFU-4 polypeptide
US-09-340-736E-10

  Query Match             57.6%;  Score 60.5;  DB 4;  Length 118;
  Best Local Similarity   58.6%;  Pred. No. 0.33;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db         38 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 66


RESULT 12
US-09-964-662-10
; Sequence 10, Application US/09964662
; Patent No. 6765086
; GENERAL INFORMATION:
;  APPLICANT: PROTEIN SPECIALTIES LTD.
;  APPLICANT:  HSC RESEARCH AND DEVELOPMENT LIMITED PARTNERSHIP
;  TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN ELASTIN AND
;  TITLE OF INVENTION:  OTHER FIBROUS PROTEINS
;  FILE REFERENCE: 041082/0112
;  CURRENT APPLICATION NUMBER: US/09/964,662
;  CURRENT FILING DATE:  2003-05-08
;  PRIOR APPLICATION NUMBER: 09/340,736
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 118
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-964-662-10

  Query Match             57.6%;  Score 60.5;  DB 4;  Length 118;
  Best Local Similarity   58.6%;  Pred. No. 0.33;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db         38 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 66


RESULT 13
US-09-340-736E-11
; Sequence 11, Application US/09340736E
; Patent No. 6489446
; GENERAL INFORMATION:
;  APPLICANT: ROTHSTEIN, ASER
;  APPLICANT:  KEELEY, FRED
;  APPLICANT:  ROTHSTEIN, STEVEN
;  TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN ELASTIN
;  TITLE OF INVENTION:  AND OTHER FIBROUS PROTEINS
;  FILE REFERENCE: 041082/0110
;  CURRENT APPLICATION NUMBER: US/09/340,736E
;  CURRENT FILING DATE:  1999-06-29
;  PRIOR APPLICATION NUMBER: 08/911,364
;  PRIOR FILING DATE: 1997-08-07
;  PRIOR APPLICATION NUMBER: 60/023,552
;  PRIOR FILING DATE: 1996-08-07
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 199
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: MFU-5 polypeptide
US-09-340-736E-11

  Query Match             57.6%;  Score 60.5;  DB 4;  Length 199;
  Best Local Similarity   58.6%;  Pred. No. 0.59;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
```

```
Db         37 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 65


RESULT 14
US-09-964-662-11
; Sequence 11, Application US/09964662
; Patent No. 6765086
; GENERAL INFORMATION:
;  APPLICANT: PROTEIN SPECIALTIES LTD.
;  APPLICANT:  HSC RESEARCH AND DEVELOPMENT LIMITED PARTNERSHIP
;  TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN ELASTIN AND
;  TITLE OF INVENTION:  OTHER FIBROUS PROTEINS
;  FILE REFERENCE: 041082/0112
;  CURRENT APPLICATION NUMBER: US/09/964,662
;  CURRENT FILING DATE:  2003-05-08
;  PRIOR APPLICATION NUMBER: 09/340,736
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 199
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-964-662-11

  Query Match             57.6%;  Score 60.5;  DB 4;  Length 199;
  Best Local Similarity   58.6%;  Pred. No. 0.59;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db         37 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 65


RESULT 15
US-09-340-736E-2
; Sequence 2, Application US/09340736E
; Patent No. 6489446
; GENERAL INFORMATION:
;  APPLICANT: ROTHSTEIN, ASER
;  APPLICANT:  KEELEY, FRED
;  APPLICANT:  ROTHSTEIN, STEVEN
;  TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN ELASTIN
;  TITLE OF INVENTION:  AND OTHER FIBROUS PROTEINS
;  FILE REFERENCE: 041082/0110
;  CURRENT APPLICATION NUMBER: US/09/340,736E
;  CURRENT FILING DATE:  1999-06-29
;  PRIOR APPLICATION NUMBER: 08/911,364
;  PRIOR FILING DATE: 1997-08-07
;  PRIOR APPLICATION NUMBER: 60/023,552
;  PRIOR FILING DATE: 1996-08-07
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 200
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: MFU-1 polypeptide
US-09-340-736E-2

  Query Match             57.6%;  Score 60.5;  DB 4;  Length 200;
  Best Local Similarity   58.6%;  Pred. No. 0.59;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db         38 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 66


RESULT 16
US-09-964-662-2
; Sequence 2, Application US/09964662
; Patent No. 6765086
; GENERAL INFORMATION:
;  APPLICANT: PROTEIN SPECIALTIES LTD.
;  APPLICANT:  HSC RESEARCH AND DEVELOPMENT LIMITED PARTNERSHIP
;  TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN ELASTIN AND
;  TITLE OF INVENTION:  OTHER FIBROUS PROTEINS
;  FILE REFERENCE: 041082/0112
;  CURRENT APPLICATION NUMBER: US/09/964,662
;  CURRENT FILING DATE:  2003-05-08
;  PRIOR APPLICATION NUMBER: 09/340,736
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 200
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: polypeptide
US-09-964-662-2

  Query Match             57.6%;  Score 60.5;  DB 4;  Length 200;
  Best Local Similarity   58.6%;  Pred. No. 0.59;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db         38 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 66


RESULT 17
US-08-911-364-2
; Sequence 2, Application US/08911364
; Patent No. 5969106
;  GENERAL INFORMATION:
;    APPLICANT:  ROTHSTEIN, Aser
;    APPLICANT:  KEELY, Fred W.
;    APPLICANT:  ROTHSTEIN, Steven J.
;    TITLE OF INVENTION:  SELF-ALIGNING PEPTIDES MODELED ON HUMAN
;    TITLE OF INVENTION:  ELASTIN AND OTHER FIBROUS PROTEINS
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY & LARDNER
;      STREET:  3000 K Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/911,364
;      FILING DATE:  07-AUG-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/023,552
;      FILING DATE:  07-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  041082/0104
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
;      TELEFAX:  (202) 672-5399
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  201 amino acids
;      TYPE:  amino acid
```

```
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-911-364-2

  Query Match             57.6%;  Score 60.5;  DB 2;  Length 201;
  Best Local Similarity   58.6%;  Pred. No. 0.6;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db         38 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 66


RESULT 18
US-09-961-403-8
; Sequence 8, Application US/09961403
; Patent No. 6780594
; GENERAL INFORMATION:
;  APPLICANT: HE-STUMPP, HOLGER
;  APPLICANT:  HAENDLER, BERNARD
;  APPLICANT:  KRAETZSCHMAR, JOERN
;  APPLICANT:  KREFT, BERTHOLT
;  APPLICANT:  WINTERHAGER, ELKE
;  APPLICANT:  REGIDOR, PEDRO
;  APPLICANT:  SCOTTI, SIMONE
;  TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
;  FILE REFERENCE: SCH-1789
;  CURRENT APPLICATION NUMBER: US/09/961,403
;  CURRENT FILING DATE:  2001-09-25
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 730
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-961-403-8

  Query Match             57.6%;  Score 60.5;  DB 4;  Length 730;
  Best Local Similarity   58.6%;  Pred. No. 2.5;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db        441 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 469


RESULT 19
US-08-911-364-1
; Sequence 1, Application US/08911364
; Patent No. 5969106
;  GENERAL INFORMATION:
;    APPLICANT:  ROTHSTEIN, Aser
;    APPLICANT:  KEELY, Fred W.
;    APPLICANT:  ROTHSTEIN, Steven J.
;    TITLE OF INVENTION:  SELF-ALIGNING PEPTIDES MODELED ON HUMAN
;    TITLE OF INVENTION:  ELASTIN AND OTHER FIBROUS PROTEINS
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FOLEY & LARDNER
;      STREET:  3000 K Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/911,364
;      FILING DATE:  07-AUG-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/023,552
;      FILING DATE:  07-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  041082/0104
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 672-5300
;      TELEFAX:  (202) 672-5399
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  731 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-911-364-1

  Query Match             57.6%;  Score 60.5;  DB 2;  Length 731;
  Best Local Similarity   58.6%;  Pred. No. 2.5;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db        415 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 443


RESULT 20
US-09-340-736E-1
; Sequence 1, Application US/09340736E
; Patent No. 6489446
; GENERAL INFORMATION:
;  APPLICANT: ROTHSTEIN, ASER
;  APPLICANT:  KEELEY, FRED
;  APPLICANT:  ROTHSTEIN, STEVEN
;  TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN ELASTIN
;  TITLE OF INVENTION:  AND OTHER FIBROUS PROTEINS
;  FILE REFERENCE: 041082/0110
;  CURRENT APPLICATION NUMBER: US/09/340,736E
;  CURRENT FILING DATE:  1999-06-29
;  PRIOR APPLICATION NUMBER: 08/911,364
;  PRIOR FILING DATE: 1997-08-07
;  PRIOR APPLICATION NUMBER: 60/023,552
;  PRIOR FILING DATE: 1996-08-07
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 731
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-340-736E-1

  Query Match             57.6%;  Score 60.5;  DB 4;  Length 731;
  Best Local Similarity   58.6%;  Pred. No. 2.5;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db        415 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 443


RESULT 21
US-09-964-662-1
; Sequence 1, Application US/09964662
; Patent No. 6765086
; GENERAL INFORMATION:
;  APPLICANT: PROTEIN SPECIALTIES LTD.
;  APPLICANT:  HSC RESEARCH AND DEVELOPMENT LIMITED PARTNERSHIP
;  TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN ELASTIN AND
;  TITLE OF INVENTION:  OTHER FIBROUS PROTEINS
;  FILE REFERENCE: 041082/0112
```

```
;  CURRENT APPLICATION NUMBER: US/09/964,662
;  CURRENT FILING DATE:  2003-05-08
;  PRIOR APPLICATION NUMBER: 09/340,736
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 731
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-964-662-1

  Query Match             57.6%;  Score 60.5;  DB 4;  Length 731;
  Best Local Similarity   58.6%;  Pred. No. 2.5;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db        415 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 443


RESULT 22
US-08-464-700-2
; Sequence 2, Application US/08464700
; Patent No. 6232458
;  GENERAL INFORMATION:
;    APPLICANT:  WEISS, ANTHONY S
;    APPLICANT:  MARTIN, STEPHEN L
;    TITLE OF INVENTION:  SYNTHETIC POLYNUCLEOTIDES
;    NUMBER OF SEQUENCES:  54
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/464,700
;      FILING DATE:  7-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PL6520
;      FILING DATE:  22-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PL9661
;      FILING DATE:  28-JUN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/AU93/00655
;      FILING DATE:  16-DEC-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  GHC3USA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  733 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-464-700-2

  Query Match             57.6%;  Score 60.5;  DB 3;  Length 733;
  Best Local Similarity   58.6%;  Pred. No. 2.6;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAAKY-----AAEAAEKAAKAX 24
              |:| || |||||     || ||: ||||:
Db        417 AQAAAAAKAAKYGVGTPAAAAAKAAAKAA 445


RESULT 23
US-09-869-875-7
; Sequence 7, Application US/09869875
; Patent No. 6521456
; GENERAL INFORMATION:
;  APPLICANT: Siebenkotten, Gregor
;  APPLICANT:  Christine, Rainer
;  TITLE OF INVENTION: USE OF CELLULAR TRANSPORT SYSTEMS FOR THE TRANSFER OF NUCLEIC ACI
;  TITLE OF INVENTION:  THROUGH THE NUCLEAR ENVELOPE
;  FILE REFERENCE: 30430.1USWO
;  CURRENT APPLICATION NUMBER: US/09/869,875
;  CURRENT FILING DATE:  2001-07-06
;  PRIOR APPLICATION NUMBER: PCT/DE00/00061
;  PRIOR FILING DATE: 2000-01-03
;  PRIOR APPLICATION NUMBER: DE 199 00 513.3
;  PRIOR FILING DATE: 1999-01-08
;  PRIOR APPLICATION NUMBER: DE 199 33 939.2
;  PRIOR FILING DATE: 1999-07-20
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
;   LENGTH: 67
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PNA-NLS
US-09-869-875-7

  Query Match             56.7%;  Score 59.5;  DB 4;  Length 67;
  Best Local Similarity   61.5%;  Pred. No. 0.24;
  Matches   16;  Conservative    6;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 AXAEAAEKAAKYAA-EAAEKAAKAXA 25
              |: ||||:||: || ||||:||: :|
Db          4 AAEEAAEEAAEEAAEEAAEEAAEEAA 29


RESULT 24
US-09-405-743A-2
; Sequence 2, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
;  APPLICANT:  Yeda Research and Development Co., Ltd.
;  TITLE OF INVENTION:  GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
;  FILE REFERENCE:  60807-A
;  CURRENT APPLICATION NUMBER: US/09/405,743A
;  CURRENT FILING DATE:  1999-09-24
;  NUMBER OF SEQ ID NOS:  7
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 45
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  SYNTHETIC
;   OTHER INFORMATION: PEPTIDE
US-09-405-743A-2

  Query Match             56.2%;  Score 59;  DB 4;  Length 45;
  Best Local Similarity   58.3%;  Pred. No. 0.18;
  Matches   14;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :|  |:|||||   ||||||  :|
Db         18 KAAEAKKAAKYEKAAAEKAAAKEA 41
```

```
RESULT 25
US-09-816-989A-2
; Sequence 2, Application US/09816989A
; Patent No. 6800287
; GENERAL INFORMATION:
;  APPLICANT: Gad, Alexander
;  APPLICANT:  Lis, Doris
;  TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
;  TITLE OF INVENTION:  AND FOR THERAPEUTIC USE
;  FILE REFERENCE: 2609/60807-A-PCT-US
;  CURRENT APPLICATION NUMBER: US/09/816,989A
;  CURRENT FILING DATE:  2001-03-23
;  PRIOR APPLICATION NUMBER: 60/101,693
;  PRIOR FILING DATE: 1998-09-25
;  PRIOR APPLICATION NUMBER: PCT/US99/22402
;  PRIOR FILING DATE: 1999-09-24
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 45
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-2

  Query Match             56.2%;  Score 59;  DB 4;  Length 45;
  Best Local Similarity   58.3%;  Pred. No. 0.18;
  Matches   14;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :|  |:|||||   ||||||  :|
Db         18 KAAEAKKAAKYEKAAAEKAAAKEA 41


RESULT 26
US-09-405-743A-7
; Sequence 7, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
;  APPLICANT:  Yeda Research and Development Co., Ltd.
;  TITLE OF INVENTION:  GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
;  FILE REFERENCE:  60807-A
;  CURRENT APPLICATION NUMBER: US/09/405,743A
;  CURRENT FILING DATE:  1999-09-24
;  NUMBER OF SEQ ID NOS:  7
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 109
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:   SYNTHETIC
;   OTHER INFORMATION: PEPTIDE
US-09-405-743A-7

  Query Match             56.2%;  Score 59;  DB 4;  Length 109;
  Best Local Similarity   56.0%;  Pred. No. 0.47;
  Matches   14;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|: || |  | ||||: ||| :|
Db         82 AEAKKAEAAKAYKAEAAKAAAKEAA 106


RESULT 27
US-09-816-989A-7
; Sequence 7, Application US/09816989A
; Patent No. 6800287
; GENERAL INFORMATION:
;  APPLICANT: Gad, Alexander
;  APPLICANT:  Lis, Doris
;  TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
;  TITLE OF INVENTION:  AND FOR THERAPEUTIC USE
;  FILE REFERENCE: 2609/60807-A-PCT-US
;  CURRENT APPLICATION NUMBER: US/09/816,989A
;  CURRENT FILING DATE:  2001-03-23
;  PRIOR APPLICATION NUMBER: 60/101,693
;  PRIOR FILING DATE: 1998-09-25
;  PRIOR APPLICATION NUMBER: PCT/US99/22402
;  PRIOR FILING DATE: 1999-09-24
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
;   LENGTH: 109
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-7

  Query Match             56.2%;  Score 59;  DB 4;  Length 109;
  Best Local Similarity   56.0%;  Pred. No. 0.47;
  Matches   14;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|: || |  | ||||: ||| :|
Db         82 AEAKKAEAAKAYKAEAAKAAAKEAA 106


RESULT 28
US-09-405-743A-5
; Sequence 5, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
;  APPLICANT:  Yeda Research and Development Co., Ltd.
;  TITLE OF INVENTION:  GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
;  FILE REFERENCE:  60807-A
;  CURRENT APPLICATION NUMBER: US/09/405,743A
;  CURRENT FILING DATE:  1999-09-24
;  NUMBER OF SEQ ID NOS:  7
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 77
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:   SYNTHETIC
;   OTHER INFORMATION: PEPTIDE
US-09-405-743A-5

  Query Match             55.2%;  Score 58;  DB 4;  Length 77;
  Best Local Similarity   51.6%;  Pred. No. 0.44;
  Matches   16;  Conservative    4;  Mismatches    5;  Indels    6;  Gaps    1;

Qy          1 AXAEAAEKAAK------YAAEAAEKAAKAXA 25
              |:|: ||||||      | |  |:| |||:|
Db         10 AYAKKAEKAAKKAEAKAYKAAEAKKKAKAEA 40


RESULT 29
US-09-816-989A-5
; Sequence 5, Application US/09816989A
; Patent No. 6800287
; GENERAL INFORMATION:
;  APPLICANT: Gad, Alexander
;  APPLICANT:  Lis, Doris
;  TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARKI
;  TITLE OF INVENTION:  AND FOR THERAPEUTIC USE
;  FILE REFERENCE: 2609/60807-A-PCT-US
;  CURRENT APPLICATION NUMBER: US/09/816,989A
;  CURRENT FILING DATE:  2001-03-23
;  PRIOR APPLICATION NUMBER: 60/101,693
;  PRIOR FILING DATE: 1998-09-25
;  PRIOR APPLICATION NUMBER: PCT/US99/22402
;  PRIOR FILING DATE: 1999-09-24
```

```
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 77
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-5

  Query Match             55.2%;  Score 58;  DB 4;  Length 77;
  Best Local Similarity   51.6%;  Pred. No. 0.44;
  Matches   16;  Conservative    4;  Mismatches    5;  Indels    6;  Gaps    1;

Qy          1 AXAEAAEKAAK------YAAEAAEKAAKAXA 25
              |:|: ||||||      | |  |:| |||:|
Db         10 AYAKKAEKAAKKAEAKAYKAAEAKKKAKAEA 40


RESULT 30
US-09-344-529-2
; Sequence 2, Application US/09344529
; Patent No. 6429293
; GENERAL INFORMATION:
;  APPLICANT: Hew, Choy L.
;  APPLICANT: HSC Research and Development Limited Partnership
;  TITLE OF INVENTION: Sculpin-Type Antifreeze Polypeptides and Nucleic Acids
;  FILE REFERENCE: 016252-002620US
;  CURRENT APPLICATION NUMBER: US/09/344,529
;  CURRENT FILING DATE: 1999-06-24
;  EARLIER APPLICATION NUMBER: US 60/090,794
;  EARLIER FILING DATE: 1998-06-26
;  EARLIER APPLICATION NUMBER: US 60/095,713
;  EARLIER FILING DATE: 1998-08-07
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 92
;   TYPE: PRT
;   ORGANISM: Myoxocephalus scorpius
US-09-344-529-2

  Query Match             55.2%;  Score 58;  DB 4;  Length 92;
  Best Local Similarity   60.0%;  Pred. No. 0.53;
  Matches   15;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|:||| ||  || ||| ||  :|
Db          3 AAAKAAEAAAMAAANAAEAAATKAA 27


RESULT 31
US-09-902-540-13678
; Sequence 13678, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13678
;   LENGTH: 325
;   TYPE: PRT
;   ORGANISM: Myxococcus xanthus
US-09-902-540-13678

  Query Match             54.8%;  Score 57.5;  DB 4;  Length 325;
  Best Local Similarity   61.5%;  Pred. No. 2.6;
  Matches   16;  Conservative    4;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAE-AAEKAAKAXA 25
              |:  ||| ||| ||| ||:| |:|:|
Db         20 AAKRAAEAAAKKAAEAAAKKTAEAAA 45


RESULT 32
US-08-303-025-12
; Sequence 12, Application US/08303025
; Patent No. 5614494
; GENERAL INFORMATION:
;    APPLICANT:  Wakefield, Thomas W.
;    APPLICANT:  Andrews, Philip C.
;    APPLICANT:  Stanley, James C.
;    TITLE OF INVENTION:  NOVEL PEPTIDES FOR HEPARIN AND
;    TITLE OF INVENTION:  LOW MOLECULAR WEIGHT HEPARIN
;    TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benita J, Rohm, Esq.
;      STREET:  150 West Jefferson, Suite 2500
;      CITY:  Detroit
;      STATE:  Michigan
;      COUNTRY:  United States of America
;      ZIP:  48226-4415
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy diskette 3.5"  1.44Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS v.6.22
;      SOFTWARE:  WordPerfect  6.1; ASCII (DOS)Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/303,025
;      FILING DATE:  08-SEPT-1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06829
;      FILING DATE:  14-AUG-1992
;      APPLICATION NUMBER:  US 08/152,488
;      FILING DATE:  12-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rohm, Benita J.
;      REFERENCE/DOCKET NUMBER:  7WH-060548-00231
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  313-496-7622
;      TELEFAX:  313-496-8454
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  28 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  N/A
;      TOPOLOGY:  N/A
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  N/A
;    PUBLICATION INFORMATION:
;      AUTHORS:  N/A
;      TITLE:  N/A
;      DOCUMENT NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
US-08-303-025-12

  Query Match             54.3%;  Score 57;  DB 1;  Length 28;
  Best Local Similarity   56.5%;  Pred. No. 0.19;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|: |:|||| | :||:|| ||:
Db          1 PAKKAKKAAKKAKKAAKKAKKAA 23
```

```
RESULT 33
US-08-436-703B-1
; Sequence 1, Application US/08436703B
; Patent No. 5919761
;  GENERAL INFORMATION:
;    APPLICANT:  Wakefield, Thomas W.
;    APPLICANT:  Andrews, Philip C.
;    APPLICANT:  Stanley, James C.
;    TITLE OF INVENTION:  NOVEL PEPTIDES FOR
;    TITLE OF INVENTION:  HEPARIN AND LOW MOLECULAR
;    TITLE OF INVENTION:  WEIGHT HEPARIN
;    TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benita J, Rohm, Esq.
;      STREET:  6601 Woodward Avenue
;      STREET:  Suite 1525
;      CITY:  Detroit
;      STATE:  Michigan
;      COUNTRY:  United States of America
;      ZIP:  48226
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk 1.44Mb, 3.5"
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 6;
;      SOFTWARE:  ASCII (DOS)Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/436,703B
;      FILING DATE:  08-MAY-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  N/A
;      FILING DATE:  N/A
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rohm, Benita J.
;      REGISTRATION NUMBER:  28,664
;      REFERENCE/DOCKET NUMBER:  7WK-060548-00233
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  313-965-1976
;      TELEFAX:  313-965-1951
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  28 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  N/A
;      TOPOLOGY:  N/A
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  N/A
;    PUBLICATION INFORMATION:
;      AUTHORS:  N/A
;      TITLE:  N/A
US-08-436-703B-1

  Query Match             54.3%;  Score 57;  DB 2;  Length 28;
  Best Local Similarity   56.5%;  Pred. No. 0.19;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|: |:|||| | :||:|| ||:
Db          1 PAKKAKKAAKKAKKAAKKAKKAA 23


RESULT 34
US-08-152-488-10
; Sequence 10, Application US/08152488
; Patent No. 5534619
;  GENERAL INFORMATION:
;    APPLICANT:  Wakefield, Thomas W.
;    APPLICANT:  Andrews, Philip C.
;    APPLICANT:  Stanley, James C.
;    TITLE OF INVENTION:  NOVEL PEPTIDES FOR HEPARIN AND
;    TITLE OF INVENTION:  LOW MOLECULAR WEIGHT HEPARIN
;    TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benita J, Rohm, Esq.
;      STREET:  512 Springfield Avenue
;      CITY:  Cranford
;      STATE:  New Jersey
;      COUNTRY:  United States of America
;      ZIP:  07016-1811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect  6; ASCII (DOS)Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/152,488
;      FILING DATE:  12-NOV-1993
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rohm, Benita J.
;      REGISTRATION NUMBER:  28,664
;      REFERENCE/DOCKET NUMBER:  RM-7WG
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-276-3344
;      TELEFAX:  908-276-5543
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  N/A
;      TOPOLOGY:  N/A
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  N/A
;    PUBLICATION INFORMATION:
;      AUTHORS:  N/A
;      TITLE:  N/A
;    PUBLICATION INFORMATION:
;      DOCUMENT NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
US-08-152-488-10

  Query Match             54.3%;  Score 57;  DB 1;  Length 29;
  Best Local Similarity   56.5%;  Pred. No. 0.2;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|: |:|||| | :||:|| ||:
Db          5 AAKKAKKAAKKAKKAAKKAKKAA 27


RESULT 35
US-08-152-488-11
; Sequence 11, Application US/08152488
; Patent No. 5534619
;  GENERAL INFORMATION:
;    APPLICANT:  Wakefield, Thomas W.
;    APPLICANT:  Andrews, Philip C.
;    APPLICANT:  Stanley, James C.
;    TITLE OF INVENTION:  NOVEL PEPTIDES FOR HEPARIN AND
;    TITLE OF INVENTION:  LOW MOLECULAR WEIGHT HEPARIN
;    TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benita J, Rohm, Esq.
;      STREET:  512 Springfield Avenue
;      CITY:  Cranford
;      STATE:  New Jersey
;      COUNTRY:  United States of America
```

```
;       ZIP:  07016-1811
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect  6; ASCII (DOS)Text
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/152,488
;       FILING DATE:  12-NOV-1993
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US92/08069
;       FILING DATE:  14-AUG-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Rohm, Benita J.
;       REGISTRATION NUMBER:  28,664
;       REFERENCE/DOCKET NUMBER:  RM-7WG
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  908-276-3344
;       TELEFAX:  908-276-5543
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  29 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  N/A
;       TOPOLOGY:  N/A
;     MOLECULE TYPE:  peptide
;     ORIGINAL SOURCE:
;       ORGANISM:  N/A
;     PUBLICATION INFORMATION:
;       AUTHORS:  N/A
;       TITLE:  N/A
;     PUBLICATION INFORMATION:
;       DOCUMENT NUMBER:  PCT/US92/08069
;       FILING DATE:  14-AUG-1993
US-08-152-488-11

  Query Match           54.3%;  Score 57;  DB 1;  Length 29;
  Best Local Similarity   56.5%;  Pred. No. 0.2;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|: |:|||| | :||:|| ||:
Db          5 AAKKAKKAAKKAKKAAKKAKKAA 27


RESULT 36
US-08-152-488-12
; Sequence 12, Application US/08152488
; Patent No. 5534619
;  GENERAL INFORMATION:
;    APPLICANT:  Wakefield, Thomas W.
;    APPLICANT:  Andrews, Philip C.
;    APPLICANT:  Stanley, James C.
;    TITLE OF INVENTION:  NOVEL PEPTIDES FOR HEPARIN AND
;    TITLE OF INVENTION:  LOW MOLECULAR WEIGHT HEPARIN
;    TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benita J, Rohm, Esq.
;      STREET:  512 Springfield Avenue
;      CITY:  Cranford
;      STATE:  New Jersey
;      COUNTRY:  United States of America
;      ZIP:  07016-1811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect  6; ASCII (DOS)Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/152,488
;      FILING DATE:  12-NOV-1993
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rohm, Benita J.
;      REGISTRATION NUMBER:  28,664
;      REFERENCE/DOCKET NUMBER:  RM-7WG
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-276-3344
;      TELEFAX:  908-276-5543
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  N/A
;      TOPOLOGY:  N/A
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  N/A
;    PUBLICATION INFORMATION:
;      AUTHORS:  N/A
;      TITLE:  N/A
;    PUBLICATION INFORMATION:
;      DOCUMENT NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
US-08-152-488-12

  Query Match           54.3%;  Score 57;  DB 1;  Length 29;
  Best Local Similarity   56.5%;  Pred. No. 0.2;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|: |:|||| | :||:|| ||:
Db          5 AAKKAKKAAKKAKKAAKKAKKAK 27


RESULT 37
US-08-303-025-10
; Sequence 10, Application US/08303025
; Patent No. 5614494
;  GENERAL INFORMATION:
;    APPLICANT:  Wakefield, Thomas W.
;    APPLICANT:  Andrews, Philip C.
;    APPLICANT:  Stanley, James C.
;    TITLE OF INVENTION:  NOVEL PEPTIDES FOR HEPARIN AND
;    TITLE OF INVENTION:  LOW MOLECULAR WEIGHT HEPARIN
;    TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benita J, Rohm, Esq.
;      STREET:  150 West Jefferson, Suite 2500
;      CITY:  Detroit
;      STATE:  Michigan
;      COUNTRY:  United States of America
;      ZIP:  48226-4415
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy diskette 3.5"  1.44Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS v.6.22
;      SOFTWARE:  WordPerfect  6.1; ASCII (DOS)Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/303,025
;      FILING DATE:  08-SEPT-1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06829
;      FILING DATE:  14-AUG-1992
;      APPLICATION NUMBER:  US 08/152,488
;      FILING DATE:  12-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rohm, Benita J.
;      REFERENCE/DOCKET NUMBER:  7WH-060548-00231
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  313-496-7622
;      TELEFAX:  313-496-8454
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  N/A
;      TOPOLOGY:  N/A
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  N/A
;    PUBLICATION INFORMATION:
;      AUTHORS:  N/A
;      TITLE:  N/A
;      DOCUMENT NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
US-08-303-025-10

  Query Match             54.3%;  Score 57;  DB 1;  Length 29;
  Best Local Similarity   56.5%;  Pred. No. 0.2;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|: |:|||| | :||:|| ||:
Db          5 AAKKAKKAAKKAKKAAKKAKKAA 27


RESULT 38
US-08-303-025-11
; Sequence 11, Application US/08303025
; Patent No. 5614494
;  GENERAL INFORMATION:
;    APPLICANT:  Wakefield, Thomas W.
;    APPLICANT:  Andrews, Philip C.
;    APPLICANT:  Stanley, James C.
;    TITLE OF INVENTION:  NOVEL PEPTIDES FOR HEPARIN AND
;    TITLE OF INVENTION:  LOW MOLECULAR WEIGHT HEPARIN
;    TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benita J, Rohm, Esq.
;      STREET:  150 West Jefferson, Suite 2500
;      CITY:  Detroit
;      STATE:  Michigan
;      COUNTRY:  United States of America
;      ZIP:  48226-4415
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy diskette 3.5"  1.44Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS v.6.22
;      SOFTWARE:  WordPerfect  6.1; ASCII (DOS)Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/303,025
;      FILING DATE:  08-SEPT-1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06829
;      FILING DATE:  14-AUG-1992
;      APPLICATION NUMBER:  US 08/152,488
;      FILING DATE:  12-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rohm, Benita J.
;      REFERENCE/DOCKET NUMBER:  7WH-060548-00231
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  313-496-7622
;      TELEFAX:  313-496-8454
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  N/A
;      TOPOLOGY:  N/A
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  N/A
;    PUBLICATION INFORMATION:
;      AUTHORS:  N/A
;      TITLE:  N/A
;      DOCUMENT NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
US-08-303-025-11

  Query Match             54.3%;  Score 57;  DB 1;  Length 29;
  Best Local Similarity   56.5%;  Pred. No. 0.2;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|: |:|||| | :||:|| ||:
Db          5 AAKKAKKAAKKAKKAAKKAKKAA 27


RESULT 39
US-08-303-025-13
; Sequence 13, Application US/08303025
; Patent No. 5614494
;  GENERAL INFORMATION:
;    APPLICANT:  Wakefield, Thomas W.
;    APPLICANT:  Andrews, Philip C.
;    APPLICANT:  Stanley, James C.
;    TITLE OF INVENTION:  NOVEL PEPTIDES FOR HEPARIN AND
;    TITLE OF INVENTION:  LOW MOLECULAR WEIGHT HEPARIN
;    TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benita J, Rohm, Esq.
;      STREET:  150 West Jefferson, Suite 2500
;      CITY:  Detroit
;      STATE:  Michigan
;      COUNTRY:  United States of America
;      ZIP:  48226-4415
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy diskette 3.5"  1.44Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS v.6.22
;      SOFTWARE:  WordPerfect  6.1; ASCII (DOS)Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/303,025
;      FILING DATE:  08-SEPT-1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06829
;      FILING DATE:  14-AUG-1992
;      APPLICATION NUMBER:  US 08/152,488
;      FILING DATE:  12-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rohm, Benita J.
;      REFERENCE/DOCKET NUMBER:  7WH-060548-00231
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  313-496-7622
;      TELEFAX:  313-496-8454
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  N/A
;      TOPOLOGY:  N/A
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  N/A
;    PUBLICATION INFORMATION:
;      AUTHORS:  N/A
;      TITLE:  N/A
;      DOCUMENT NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
US-08-303-025-13
```

```
  Query Match             54.3%;  Score 57;  DB 1;  Length 29;
  Best Local Similarity   56.5%;  Pred. No. 0.2;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|: |:|||| | :||:|| ||:
Db          5 AAKKAKKAAKKAKKAAKKAKKAA 27


RESULT 40
US-08-303-025-14
; Sequence 14, Application US/08303025
; Patent No. 5614494
;  GENERAL INFORMATION:
;    APPLICANT:  Wakefield, Thomas W.
;    APPLICANT:  Andrews, Philip C.
;    APPLICANT:  Stanley, James C.
;    TITLE OF INVENTION:  NOVEL PEPTIDES FOR HEPARIN AND
;    TITLE OF INVENTION:  LOW MOLECULAR WEIGHT HEPARIN
;    TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benita J, Rohm, Esq.
;      STREET:  150 West Jefferson, Suite 2500
;      CITY:  Detroit
;      STATE:  Michigan
;      COUNTRY:  United States of America
;      ZIP:  48226-4415
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: /Floppy diskette 3.5"  1.44Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS v.6.22
;      SOFTWARE:  WordPerfect  6.1; ASCII (DOS)Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/303,025
;      FILING DATE:  08-SEPT-1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06829
;      FILING DATE:  14-AUG-1992
;      APPLICATION NUMBER:  US 08/152,488
;      FILING DATE:  12-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rohm, Benita J.
;      REFERENCE/DOCKET NUMBER:  7WH-060548-00231
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  313-496-7622
;      TELEFAX:  313-496-8454
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  N/A
;      TOPOLOGY:  N/A
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  N/A
;    PUBLICATION INFORMATION:
;      AUTHORS:  N/A
;      TITLE:  N/A
;      DOCUMENT NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
US-08-303-025-14

  Query Match             54.3%;  Score 57;  DB 1;  Length 29;
  Best Local Similarity   56.5%;  Pred. No. 0.2;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|: |:|||| | :||:|| ||:
Db          5 AAKKAKKAAKKAKKAAKKAKKAK 27


RESULT 41
US-08-677-304-10
; Sequence 10, Application US/08677304
; Patent No. 5721212
;  GENERAL INFORMATION:
;    APPLICANT:  Wakefield, Thomas W.
;    APPLICANT:  Andrews, Philip C.
;    APPLICANT:  Stanley, James C.
;    TITLE OF INVENTION:  NOVEL PEPTIDES FOR HEPARIN AND
;    TITLE OF INVENTION:  LOW MOLECULAR WEIGHT HEPARIN
;    TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benita J, Rohm, Esq.
;      STREET:  512 Springfield Avenue
;      CITY:  Cranford
;      STATE:  New Jersey
;      COUNTRY:  United States of America
;      ZIP:  07016-1811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect  6; ASCII (DOS)Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/677,304
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/152,488
;      FILING DATE:  12-NOV-1993
;      APPLICATION NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rohm, Benita J.
;      REGISTRATION NUMBER:  28,664
;      REFERENCE/DOCKET NUMBER:  RM-7WG
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-276-3344
;      TELEFAX:  908-276-5543
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  No. 5721212 Relevant
;      TOPOLOGY:  No. 5721212 Relevant
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  N/A
;    PUBLICATION INFORMATION:
;      AUTHORS:  N/A
;      TITLE:  N/A
;    PUBLICATION INFORMATION:
;      DOCUMENT NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
US-08-677-304-10

  Query Match             54.3%;  Score 57;  DB 1;  Length 29;
  Best Local Similarity   56.5%;  Pred. No. 0.2;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|: |:|||| | :||:|| ||:
Db          5 AAKKAKKAAKKAKKAAKKAKKAA 27


RESULT 42
US-08-677-304-11
; Sequence 11, Application US/08677304
; Patent No. 5721212
;  GENERAL INFORMATION:
;    APPLICANT:  Wakefield, Thomas W.
```

```
;    APPLICANT:  Andrews, Philip C.
;    APPLICANT:  Stanley, James C.
;    TITLE OF INVENTION:  NOVEL PEPTIDES FOR HEPARIN AND
;    TITLE OF INVENTION:  LOW MOLECULAR WEIGHT HEPARIN
;    TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benita J, Rohm, Esq.
;      STREET:  512 Springfield Avenue
;      CITY:  Cranford
;      STATE:  New Jersey
;      COUNTRY:  United States of America
;      ZIP:  07016-1811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect  6; ASCII (DOS)Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/677,304
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/152,488
;      FILING DATE:  12-NOV-1993
;      APPLICATION NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rohm, Benita J.
;      REGISTRATION NUMBER:  28,664
;      REFERENCE/DOCKET NUMBER:  RM-7WG
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-276-3344
;      TELEFAX:  908-276-5543
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  No. 5721212 Relevant
;      TOPOLOGY:  No. 5721212 Relevant
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  N/A
;    PUBLICATION INFORMATION:
;      AUTHORS:  N/A
;      TITLE:  N/A
;    PUBLICATION INFORMATION:
;      DOCUMENT NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
US-08-677-304-11

  Query Match             54.3%;  Score 57;  DB 1;  Length 29;
  Best Local Similarity   56.5%;  Pred. No. 0.2;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|: |:|||| | :||:|| ||:
Db          5 AAKKAKKAAKKAKKAAKKAKKAA 27


RESULT 43
US-08-677-304-12
; Sequence 12, Application US/08677304
; Patent No. 5721212
;  GENERAL INFORMATION:
;    APPLICANT:  Wakefield, Thomas W.
;    APPLICANT:  Andrews, Philip C.
;    APPLICANT:  Stanley, James C.
;    TITLE OF INVENTION:  NOVEL PEPTIDES FOR HEPARIN AND
;    TITLE OF INVENTION:  LOW MOLECULAR WEIGHT HEPARIN
;    TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benita J, Rohm, Esq.
;      STREET:  512 Springfield Avenue
;      CITY:  Cranford
;      STATE:  New Jersey
;      COUNTRY:  United States of America
;      ZIP:  07016-1811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect  6; ASCII (DOS)Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/677,304
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/152,488
;      FILING DATE:  12-NOV-1993
;      APPLICATION NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rohm, Benita J.
;      REGISTRATION NUMBER:  28,664
;      REFERENCE/DOCKET NUMBER:  RM-7WG
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-276-3344
;      TELEFAX:  908-276-5543
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  No. 5721212 Relevant
;      TOPOLOGY:  No. 5721212 Relevant
;    MOLECULE TYPE:  peptide
;    ORIGINAL SOURCE:
;      ORGANISM:  N/A
;    PUBLICATION INFORMATION:
;      AUTHORS:  N/A
;      TITLE:  N/A
;    PUBLICATION INFORMATION:
;      DOCUMENT NUMBER:  PCT/US92/08069
;      FILING DATE:  14-AUG-1993
US-08-677-304-12

  Query Match             54.3%;  Score 57;  DB 1;  Length 29;
  Best Local Similarity   56.5%;  Pred. No. 0.2;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|: |:|||| | :||:|| ||:
Db          5 AAKKAKKAAKKAKKAAKKAKKAK 27


RESULT 44
US-08-436-703B-3
; Sequence 3, Application US/08436703B
; Patent No. 5919761
;  GENERAL INFORMATION:
;    APPLICANT:  Wakefield, Thomas W.
;    APPLICANT:  Andrews, Philip C.
;    APPLICANT:  Stanley, James C.
;    TITLE OF INVENTION:  NOVEL PEPTIDES FOR
;    TITLE OF INVENTION:  HEPARIN AND LOW MOLECULAR
;    TITLE OF INVENTION:  WEIGHT HEPARIN
;    TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Benita J, Rohm, Esq.
;      STREET:  6601 Woodward Avenue
;      STREET:  Suite 1525
;      CITY:  Detroit
;      STATE:  Michigan
;      COUNTRY:  United States of America
```

```
;       ZIP:  48226
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk 1.44Mb, 3.5"
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect 6;
;       SOFTWARE:  ASCII (DOS)Text
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/436,703B
;       FILING DATE:  08-MAY-1995
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  N/A
;       FILING DATE:  N/A
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Rohm, Benita J.
;       REGISTRATION NUMBER:  28,664
;       REFERENCE/DOCKET NUMBER:  7WK-060548-00233
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  313-965-1976
;       TELEFAX:  313-965-1951
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  29 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  N/A
;       TOPOLOGY:  N/A
;     MOLECULE TYPE:  peptide
;     ORIGINAL SOURCE:
;       ORGANISM:  N/A
;     PUBLICATION INFORMATION:
;       AUTHORS:  N/A
;       TITLE:  N/A
US-08-436-703B-3

  Query Match             54.3%;  Score 57;  DB 2;  Length 29;
  Best Local Similarity   56.5%;  Pred. No. 0.2;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|: |:|||| | :||:|| ||:
Db          5 AAKKAKKAAKKAKKAAKKAKKAA 27


RESULT 45
US-08-436-703B-15
; Sequence 15, Application US/08436703B
; Patent No. 5919761
;  GENERAL INFORMATION:
;     APPLICANT:  Wakefield, Thomas W.
;     APPLICANT:  Andrews, Philip C.
;     APPLICANT:  Stanley, James C.
;     TITLE OF INVENTION:  NOVEL PEPTIDES FOR
;     TITLE OF INVENTION:  HEPARIN AND LOW MOLECULAR
;     TITLE OF INVENTION:  WEIGHT HEPARIN
;     TITLE OF INVENTION:  ANTICOAGULATION REVERSAL
;     NUMBER OF SEQUENCES:  18
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Benita J, Rohm, Esq.
;       STREET:  6601 Woodward Avenue
;       STREET:  Suite 1525
;       CITY:  Detroit
;       STATE:  Michigan
;       COUNTRY:  United States of America
;       ZIP:  48226
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk 1.44Mb, 3.5"
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  WordPerfect 6;
;       SOFTWARE:  ASCII (DOS)Text
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/436,703B
;       FILING DATE:  08-MAY-1995
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  N/A
;       FILING DATE:  N/A
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Rohm, Benita J.
;       REGISTRATION NUMBER:  28,664
;       REFERENCE/DOCKET NUMBER:  7WK-060548-00233
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  313-965-1976
;       TELEFAX:  313-965-1951
;   INFORMATION FOR SEQ ID NO:  15:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  29 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  N/A
;       TOPOLOGY:  N/A
;     MOLECULE TYPE:  peptide
;     ORIGINAL SOURCE:
;       ORGANISM:  N/A
;     PUBLICATION INFORMATION:
;       AUTHORS:  N/A
;       TITLE:  N/A
US-08-436-703B-15

  Query Match             54.3%;  Score 57;  DB 2;  Length 29;
  Best Local Similarity   56.5%;  Pred. No. 0.2;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|: |:|||| | :||:|| ||:
Db          5 AAKKAKKAAKKAKKAAKKAKKAA 27


Search completed: July 11, 2005, 09:47:56
Job time : 43 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2005, 09:46:35 ; Search time 159 Seconds
(without alignments)
60.744 Million cell updates/sec

Title: SEQ1
Perfect score: 105
Sequence: 1 axaeaaekaakyaaeaaekaakaxa 25

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1726218 seqs, 386331768 residues

Total number of hits satisfying chosen parameters: 1726218

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	67	63.8	104	14	US-10-177-725-42	Sequence 42, Appl
2	67	63.8	104	14	US-10-177-725-92	Sequence 92, Appl
3	67	63.8	104	15	US-10-393-449-42	Sequence 42, Appl
4	67	63.8	104	15	US-10-393-449-92	Sequence 92, Appl
5	66	62.9	827	16	US-10-437-963-152005	Sequence 152005,
6	65.5	62.4	428	15	US-10-282-122A-55748	Sequence 55748, A
7	64	61.0	28	16	US-10-667-004-21	Sequence 21, Appl
8	64	61.0	28	16	US-10-667-004-24	Sequence 24, Appl
9	64	61.0	104	14	US-10-177-725-39	Sequence 39, Appl
10	64	61.0	104	14	US-10-177-725-40	Sequence 40, Appl
11	64	61.0	104	14	US-10-177-725-89	Sequence 89, Appl
12	64	61.0	104	14	US-10-177-725-90	Sequence 90, Appl
13	64	61.0	104	15	US-10-393-449-39	Sequence 39, Appl
14	64	61.0	104	15	US-10-393-449-40	Sequence 40, Appl
15	64	61.0	104	15	US-10-393-449-89	Sequence 89, Appl
16	64	61.0	104	15	US-10-393-449-90	Sequence 90, Appl
17	63.5	60.5	104	14	US-10-177-725-41	Sequence 41, Appl
18	63.5	60.5	104	14	US-10-177-725-91	Sequence 91, Appl
19	63.5	60.5	104	15	US-10-393-449-41	Sequence 41, Appl
20	63.5	60.5	104	15	US-10-393-449-91	Sequence 91, Appl
21	63	60.0	59	14	US-10-177-725-55	Sequence 55, Appl
22	63	60.0	59	14	US-10-177-725-105	Sequence 105, App
23	63	60.0	59	15	US-10-393-449-55	Sequence 55, Appl
24	63	60.0	59	15	US-10-393-449-105	Sequence 105, App
25	63	60.0	67	14	US-10-177-725-54	Sequence 54, Appl
26	63	60.0	67	14	US-10-177-725-104	Sequence 104, App
27	63	60.0	67	15	US-10-393-449-54	Sequence 54, Appl
28	63	60.0	67	15	US-10-393-449-104	Sequence 104, App
29	63	60.0	75	14	US-10-177-725-53	Sequence 53, Appl
30	63	60.0	75	14	US-10-177-725-103	Sequence 103, App
31	63	60.0	75	15	US-10-393-449-53	Sequence 53, Appl
32	63	60.0	75	15	US-10-393-449-103	Sequence 103, App
33	63	60.0	83	14	US-10-177-725-52	Sequence 52, Appl
34	63	60.0	83	14	US-10-177-725-102	Sequence 102, App
35	63	60.0	83	15	US-10-393-449-52	Sequence 52, Appl
36	63	60.0	83	15	US-10-393-449-102	Sequence 102, App
37	63	60.0	88	14	US-10-177-725-49	Sequence 49, Appl
38	63	60.0	88	14	US-10-177-725-99	Sequence 99, Appl
39	63	60.0	88	15	US-10-393-449-49	Sequence 49, Appl
40	63	60.0	88	15	US-10-393-449-99	Sequence 99, Appl
41	63	60.0	91	14	US-10-177-725-51	Sequence 51, Appl
42	63	60.0	91	14	US-10-177-725-101	Sequence 101, App
43	63	60.0	91	15	US-10-393-449-51	Sequence 51, Appl
44	63	60.0	91	15	US-10-393-449-101	Sequence 101, App
45	63	60.0	104	14	US-10-177-725-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1
US-10-177-725-42
; Sequence 42, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Peele, Beau R.
; TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
; FILE REFERENCE: A-66900-4/RMS/AMS
; CURRENT APPLICATION NUMBER: US/10/177,725
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 09/415,765
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 09/169,015
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-10-177-725-42

Query Match 63.8%; Score 67; DB 14; Length 104;
Best Local Similarity 68.0%; Pred. No. 0.3;
Matches 17; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||| ||| ||||| |||  :|
Db         10 AAAAAAEAAAKAAAEAAAKAAAEAA 34
```

```
RESULT 2
US-10-177-725-92
; Sequence 92, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 92
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (37)..(68)
;   OTHER INFORMATION: "Xaa" at positions 37-39, 41-43, 45-46, 48-50, 52-53, 55-57, 59-6
;   OTHER INFORMATION: 1, 63-64 and 66-68 can be any amino acid
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (37)..(68)
;   OTHER INFORMATION: "Xaa" at positions 38-40, 42-44, 46-47, 49-51, 53-54, 56-58, 60-6
;   OTHER INFORMATION: 2, 64-65, and 67-69 can be any amino acid
US-10-177-725-92

  Query Match             63.8%;  Score 67;  DB 14;  Length 104;
  Best Local Similarity   68.0%;  Pred. No. 0.3;
  Matches   17;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||| ||| ||||| |||  :|
Db         10 AAAAAAEAAAKAAAEAAAKAAAEAA 34


RESULT 3
US-10-393-449-42
; Sequence 42, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-393-449-42

  Query Match             63.8%;  Score 67;  DB 15;  Length 104;
  Best Local Similarity   68.0%;  Pred. No. 0.3;
  Matches   17;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||| ||| ||||| |||  :|
Db         10 AAAAAAEAAAKAAAEAAAKAAAEAA 34


RESULT 4
US-10-393-449-92
; Sequence 92, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT SC
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 92
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (37)..(68)
;   OTHER INFORMATION: "Xaa" at positions 37-39, 41-43, 45-46, 48-50, 52-53, 55-57, 59-6
;   OTHER INFORMATION: 1, 63-64 and 66-68 can be any amino acid
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (37)..(68)
;   OTHER INFORMATION: "Xaa" at positions 38-40, 42-44, 46-47, 49-51, 53-54, 56-58, 60-6
;   OTHER INFORMATION: 2, 64-65, and 67-69 can be any amino acid
US-10-393-449-92

  Query Match             63.8%;  Score 67;  DB 15;  Length 104;
  Best Local Similarity   68.0%;  Pred. No. 0.3;
  Matches   17;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||| ||| ||||| |||  :|
Db         10 AAAAAAEAAAKAAAEAAAKAAAEAA 34


RESULT 5
US-10-437-963-152005
; Sequence 152005, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
```

```
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 152005
;   LENGTH: 827
;   TYPE: PRT
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_52099C.1.pep
US-10-437-963-152005

  Query Match             62.9%;  Score 66;  DB 16;  Length 827;
  Best Local Similarity   54.5%;  Pred. No. 4;
  Matches   18;  Conservative    4;  Mismatches    3;  Indels    8;  Gaps    1;

Qy          1 AXAEAAEKAAKYAA--------EAAEKAAKAXA 25
              |:||| |:||| ||        || |:||||:|
Db        398 AAAEARERAAKAAAEAKERVAEEARERAAKAAA 430


RESULT 6
US-10-282-122A-55748
; Sequence 55748, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55748
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Enterobacter cloacae
US-10-282-122A-55748

  Query Match             62.4%;  Score 65.5;  DB 15;  Length 428;
  Best Local Similarity   62.1%;  Pred. No. 2.2;
  Matches   18;  Conservative    4;  Mismatches    2;  Indels    5;  Gaps    1;

Qy          1 AXAEAAEKAA-----KYAAEAAEKAAKAX 24
              |:||||:|||     | |||||:||| |:
Db        210 AEAEAAKKAAQEAEKKAAAEAAKKAAAAE 238


RESULT 7
US-10-667-004-21
; Sequence 21, Application US/10667004
; Publication No. US20040126820A1
; GENERAL INFORMATION:
;  APPLICANT: INTEL CORPORATION
;  APPLICANT:  CHAN, Selena
;  APPLICANT:  SU, Xing
;  APPLICANT:  YAMAKAWA, Mineo
;  TITLE OF INVENTION: CONTROLLED ALIGNMENT OF NANO-BARCODES ENCODING SPECIFIC INFORMATI
;  TITLE OF INVENTION:  SCANNING PROBE MICROSCOPY (SPM)
;  FILE REFERENCE: INTEL1310-1(P14240X)
;  CURRENT APPLICATION NUMBER: US/10/667,004
;  CURRENT FILING DATE:  2003-09-19
;  PRIOR APPLICATION NUMBER: US 10/251,152
;  PRIOR FILING DATE: 2002-09-20
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
;   LENGTH: 28
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic peptide
US-10-667-004-21

  Query Match             61.0%;  Score 64;  DB 16;  Length 28;
  Best Local Similarity   65.2%;  Pred. No. 0.17;
  Matches   15;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|:|| |||| ||:||: ||||:
Db          1 AAKAAAKAAKAAAKAAKAAAKAA 23


RESULT 8
US-10-667-004-24
; Sequence 24, Application US/10667004
; Publication No. US20040126820A1
; GENERAL INFORMATION:
;  APPLICANT: INTEL CORPORATION
;  APPLICANT:  CHAN, Selena
;  APPLICANT:  SU, Xing
;  APPLICANT:  YAMAKAWA, Mineo
;  TITLE OF INVENTION: CONTROLLED ALIGNMENT OF NANO-BARCODES ENCODING SPECIFIC INFORMATI
;  TITLE OF INVENTION:  SCANNING PROBE MICROSCOPY (SPM)
;  FILE REFERENCE: INTEL1310-1(P14240X)
;  CURRENT APPLICATION NUMBER: US/10/667,004
;  CURRENT FILING DATE:  2003-09-19
;  PRIOR APPLICATION NUMBER: US 10/251,152
;  PRIOR FILING DATE: 2002-09-20
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
;   LENGTH: 28
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic peptide
US-10-667-004-24

  Query Match             61.0%;  Score 64;  DB 16;  Length 28;
  Best Local Similarity   65.2%;  Pred. No. 0.17;
  Matches   15;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAX 24
              :|||| :||: |||||| ||:|:
```

```
Db          1 AAEAAAEAAEAAAEAAEAAAEAA 23


RESULT 9
US-10-177-725-39
; Sequence 39, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-177-725-39

  Query Match             61.0%;  Score 64;  DB 14;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.73;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|||| |||  || ||| ||||:|
Db          9 AAAEAAAKAA--AAAAAEAAAKAAA 31


RESULT 10
US-10-177-725-40
; Sequence 40, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-177-725-40

  Query Match             61.0%;  Score 64;  DB 14;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.73;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|||| |||  || ||| ||||:|
Db          9 AAAEAAAKAA--AAAAAEAAAKAAA 31


RESULT 11
US-10-177-725-89
; Sequence 89, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT SC
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (37)..(68)
;   OTHER INFORMATION: "Xaa" at positions 37-39, 41-43, 45-46, 48-50, 52-53, 55-57, 59-6
;   OTHER INFORMATION: 1, 63-64 and 66-68 can be any amino acid
US-10-177-725-89

  Query Match             61.0%;  Score 64;  DB 14;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.73;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|||| |||  || ||| ||||:|
Db          9 AAAEAAAKAA--AAAAAEAAAKAAA 31


RESULT 12
US-10-177-725-90
; Sequence 90, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT SC
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 90
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (37)..(68)
;   OTHER INFORMATION: "Xaa" at positions 37-39, 41-43, 45-46, 48-50, 52-53, 55-57, 59-6
;   OTHER INFORMATION: 1, 63-64 and 66-68 can be any amino acid
US-10-177-725-90

  Query Match             61.0%;  Score 64;  DB 14;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.73;
```

```
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|||| |||  || ||| ||||:|
Db          9 AAAEAAAKAA--AAAAAEAAAKAAA 31


RESULT 13
US-10-393-449-39
; Sequence 39, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-393-449-39

  Query Match             61.0%;  Score 64;  DB 15;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.73;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|||| |||  || ||| ||||:|
Db          9 AAAEAAAKAA--AAAAAEAAAKAAA 31


RESULT 14
US-10-393-449-40
; Sequence 40, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-393-449-40

  Query Match             61.0%;  Score 64;  DB 15;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.73;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|||| |||  || ||| ||||:|
Db          9 AAAEAAAKAA--AAAAAEAAAKAAA 31


RESULT 15
US-10-393-449-89
; Sequence 89, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (37)..(68)
;   OTHER INFORMATION: "Xaa" at positions 37-39, 41-43, 45-46, 48-50, 52-53, 55-57, 59-6
;   OTHER INFORMATION: 1, 63-64 and 66-68 can be any amino acid
US-10-393-449-89

  Query Match             61.0%;  Score 64;  DB 15;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.73;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|||| |||  || ||| ||||:|
Db          9 AAAEAAAKAA--AAAAAEAAAKAAA 31


RESULT 16
US-10-393-449-90
; Sequence 90, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 90
;   LENGTH: 104
;   TYPE: PRT
```

```
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (37)..(68)
;   OTHER INFORMATION: "Xaa" at positions 37-39, 41-43, 45-46, 48-50, 52-53, 55-57, 59-6
;   OTHER INFORMATION: 1, 63-64 and 66-68 can be any amino acid
US-10-393-449-90

  Query Match             61.0%;  Score 64;  DB 15;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.73;
  Matches   18;  Conservative    2;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:|||| |||  || ||| ||||:|
Db          9 AAAEAAAKAA--AAAAAEAAAKAAA 31


RESULT 17
US-10-177-725-41
; Sequence 41, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-177-725-41

  Query Match             60.5%;  Score 63.5;  DB 14;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.85;
  Matches   18;  Conservative    2;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||| ||| || ||| ||||:|
Db          6 AAAAAAEAAAK-AAAAAEAAAKAAA 29


RESULT 18
US-10-177-725-91
; Sequence 91, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 91
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (37)..(68)
;   OTHER INFORMATION: "Xaa" at positions 37-39, 41-43, 45-46, 48-50, 52-53, 55-57, 59-6
;   OTHER INFORMATION: 1, 63-64 and 66-68 can be any amino acid
US-10-177-725-91

  Query Match             60.5%;  Score 63.5;  DB 14;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.85;
  Matches   18;  Conservative    2;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||| ||| || ||| ||||:|
Db          6 AAAAAAEAAAK-AAAAAEAAAKAAA 29


RESULT 19
US-10-393-449-41
; Sequence 41, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT SC
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-393-449-41

  Query Match             60.5%;  Score 63.5;  DB 15;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.85;
  Matches   18;  Conservative    2;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||| ||| || ||| ||||:|
Db          6 AAAAAAEAAAK-AAAAAEAAAKAAA 29


RESULT 20
US-10-393-449-91
; Sequence 91, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT SC
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
```

```
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 91
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (37)..(68)
;   OTHER INFORMATION: "Xaa" at positions 37-39, 41-43, 45-46, 48-50, 52-53, 55-57, 59-6
;   OTHER INFORMATION: 1, 63-64 and 66-68 can be any amino acid
US-10-393-449-91

  Query Match             60.5%;  Score 63.5;  DB 15;  Length 104;
  Best Local Similarity   72.0%;  Pred. No. 0.85;
  Matches   18;  Conservative    2;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAXA 25
              |:| ||| ||| || ||| ||||:|
Db          6 AAAAAAEAAAK-AAAAAEAAAKAAA 29


RESULT 21
US-10-177-725-55
; Sequence 55, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
;   LENGTH: 59
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-177-725-55

  Query Match             60.0%;  Score 63;  DB 14;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 0.53;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 22
US-10-177-725-105
; Sequence 105, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 105
;   LENGTH: 59
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (30)..(46)
;   OTHER INFORMATION: "Xaa" at positions 30-32, 34-36, 38-39, 41-43, and 45-46 can be a
;   OTHER INFORMATION: ny amino acid
US-10-177-725-105

  Query Match             60.0%;  Score 63;  DB 14;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 0.53;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 23
US-10-393-449-55
; Sequence 55, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
;   LENGTH: 59
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-393-449-55

  Query Match             60.0%;  Score 63;  DB 15;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 0.53;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 24
US-10-393-449-105
; Sequence 105, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
```

```
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 105
;   LENGTH: 59
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (30)..(46)
;   OTHER INFORMATION: "Xaa" at positions 30-32, 34-36, 38-39, 41-43, and 45-46 can be a
;   OTHER INFORMATION: ny amino acid
US-10-393-449-105

  Query Match             60.0%;  Score 63;  DB 15;  Length 59;
  Best Local Similarity   66.7%;  Pred. No. 0.53;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| ||| :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 25
US-10-177-725-54
; Sequence 54, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
;   LENGTH: 67
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-177-725-54

  Query Match             60.0%;  Score 63;  DB 14;  Length 67;
  Best Local Similarity   66.7%;  Pred. No. 0.61;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| ||| :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 26
US-10-177-725-104
; Sequence 104, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT SC
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 104
;   LENGTH: 67
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (38)..(54)
;   OTHER INFORMATION: "Xaa" at positions 38-40, 42-44, 46-47, 49-51, and 53-54 can be a
;   OTHER INFORMATION: ny amino acid
US-10-177-725-104

  Query Match             60.0%;  Score 63;  DB 14;  Length 67;
  Best Local Similarity   66.7%;  Pred. No. 0.61;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| ||| :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 27
US-10-393-449-54
; Sequence 54, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT SC
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
;   LENGTH: 67
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-393-449-54

  Query Match             60.0%;  Score 63;  DB 15;  Length 67;
  Best Local Similarity   66.7%;  Pred. No. 0.61;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| ||| :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27
```

```
RESULT 28
US-10-393-449-104
; Sequence 104, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 104
;   LENGTH: 67
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (38)..(54)
;   OTHER INFORMATION: "Xaa" at positions 38-40, 42-44, 46-47, 49-51, and 53-54 can be a
;   OTHER INFORMATION: ny amino acid
US-10-393-449-104

  Query Match             60.0%;  Score 63;  DB 15;  Length 67;
  Best Local Similarity   66.7%;  Pred. No. 0.61;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 29
US-10-177-725-53
; Sequence 53, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
;   LENGTH: 75
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-177-725-53

  Query Match             60.0%;  Score 63;  DB 14;  Length 75;
  Best Local Similarity   66.7%;  Pred. No. 0.69;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 30
US-10-177-725-103
; Sequence 103, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 103
;   LENGTH: 75
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (38)..(54)
;   OTHER INFORMATION: "Xaa" at positions 38-40, 42-44, 46-47, 49-51, and 53-54 can be a
;   OTHER INFORMATION: ny amino acid
US-10-177-725-103

  Query Match             60.0%;  Score 63;  DB 14;  Length 75;
  Best Local Similarity   66.7%;  Pred. No. 0.69;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| |||  :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 31
US-10-393-449-53
; Sequence 53, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
;   LENGTH: 75
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-393-449-53
```

```
  Query Match             60.0%;  Score 63;  DB 15;  Length 75;
  Best Local Similarity   66.7%;  Pred. No. 0.69;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| ||| :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 32
US-10-393-449-103
; Sequence 103, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 103
;   LENGTH: 75
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (38)..(54)
;   OTHER INFORMATION: "Xaa" at positions 38-40, 42-44, 46-47, 49-51, and 53-54 can be a
;   OTHER INFORMATION: ny amino acid
US-10-393-449-103

  Query Match             60.0%;  Score 63;  DB 15;  Length 75;
  Best Local Similarity   66.7%;  Pred. No. 0.69;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 XAEAAEKAAKYAAEAAEKAAKAXA 25
              :| ||| ||| ||||| ||| :|
Db          4 DAAAAEAAAKAAAEAAAKAAAEAA 27


RESULT 33
US-10-177-725-52
; Sequence 52, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
;   LENGTH: 83
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-177-725-52

  Query Match             60.0%;  Score 63;  DB 14;  Length 83;
  Best Local Similarity   66.7%;  Pred. No. 0.77;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         57 AAKAAAEAAAKAAAEAAAKAAAAK 80


RESULT 34
US-10-177-725-102
; Sequence 102, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT SC
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 102
;   LENGTH: 83
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (38)..(54)
;   OTHER INFORMATION: "Xaa" at positions 38-40, 42-44, 46-47, 49-51, and 53-54 can be a
;   OTHER INFORMATION: ny amino acid
US-10-177-725-102

  Query Match             60.0%;  Score 63;  DB 14;  Length 83;
  Best Local Similarity   66.7%;  Pred. No. 0.77;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         57 AAKAAAEAAAKAAAEAAAKAAAAK 80


RESULT 35
US-10-393-449-52
; Sequence 52, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT SC
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 52
;   LENGTH: 83
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-393-449-52

  Query Match             60.0%;  Score 63;  DB 15;  Length 83;
  Best Local Similarity   66.7%;  Pred. No. 0.77;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         57 AAKAAAEAAAKAAAEAAAKAAAAK 80


RESULT 36
US-10-393-449-102
; Sequence 102, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 102
;   LENGTH: 83
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (38)..(54)
;   OTHER INFORMATION: "Xaa" at positions 38-40, 42-44, 46-47, 49-51, and 53-54 can be a
;   OTHER INFORMATION: ny amino acid
US-10-393-449-102

  Query Match             60.0%;  Score 63;  DB 15;  Length 83;
  Best Local Similarity   66.7%;  Pred. No. 0.77;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         57 AAKAAAEAAAKAAAEAAAKAAAAK 80


RESULT 37
US-10-177-725-49
; Sequence 49, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
;   LENGTH: 88
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-177-725-49

  Query Match             60.0%;  Score 63;  DB 14;  Length 88;
  Best Local Similarity   66.7%;  Pred. No. 0.82;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         63 AAKAAAEAAAKAAAEAAAKAAAAK 86


RESULT 38
US-10-177-725-99
; Sequence 99, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 99
;   LENGTH: 88
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (29)..(60)
;   OTHER INFORMATION: "Xaa" at positions 29-31, 33-35, 37-38, 40-42, 44-45, 47-49, 51-5
;   OTHER INFORMATION: 3, 55-56, and 58-60 can be any amino acid
US-10-177-725-99

  Query Match             60.0%;  Score 63;  DB 14;  Length 88;
  Best Local Similarity   66.7%;  Pred. No. 0.82;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         63 AAKAAAEAAAKAAAEAAAKAAAAK 86


RESULT 39
US-10-393-449-49
; Sequence 49, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
```

```
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
;   LENGTH: 88
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-393-449-49

  Query Match             60.0%;  Score 63;  DB 15;  Length 88;
  Best Local Similarity   66.7%;  Pred. No. 0.82;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         63 AAKAAAEAAAKAAAEAAAKAAAAK 86


RESULT 40
US-10-393-449-99
; Sequence 99, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 99
;   LENGTH: 88
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (29)..(60)
;   OTHER INFORMATION: "Xaa" at positions 29-31, 33-35, 37-38, 40-42, 44-45, 47-49, 51-5
;   OTHER INFORMATION: 3, 55-56, and 58-60 can be any amino acid
US-10-393-449-99

  Query Match             60.0%;  Score 63;  DB 15;  Length 88;
  Best Local Similarity   66.7%;  Pred. No. 0.82;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         63 AAKAAAEAAAKAAAEAAAKAAAAK 86


RESULT 41
US-10-177-725-51
; Sequence 51, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT SC
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
;   LENGTH: 91
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-177-725-51

  Query Match             60.0%;  Score 63;  DB 14;  Length 91;
  Best Local Similarity   66.7%;  Pred. No. 0.85;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         65 AAKAAAEAAAKAAAEAAAKAAAAK 88


RESULT 42
US-10-177-725-101
; Sequence 101, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT SC
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 101
;   LENGTH: 91
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (38)..(54)
;   OTHER INFORMATION: "Xaa" at positions 38-40, 42-44, 46-47, 49-51, and 53-54 can be a
;   OTHER INFORMATION: ny amino acid
US-10-177-725-101

  Query Match             60.0%;  Score 63;  DB 14;  Length 91;
  Best Local Similarity   66.7%;  Pred. No. 0.85;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         65 AAKAAAEAAAKAAAEAAAKAAAAK 88


RESULT 43
US-10-393-449-51
; Sequence 51, Application US/10393449
; Publication No. US20030224412A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
;   LENGTH: 91
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-393-449-51

  Query Match             60.0%;  Score 63;  DB 15;  Length 91;
  Best Local Similarity   66.7%;  Pred. No. 0.85;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         65 AAKAAAEAAAKAAAEAAAKAAAAK 88


RESULT 44
US-10-393-449-101
; Sequence 101, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
;  FILE REFERENCE: RIGL-007CIP3
;  CURRENT APPLICATION NUMBER: US/10/393,449
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/177,725
;  PRIOR FILING DATE: 2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 101
;   LENGTH: 91
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (38)..(54)
;   OTHER INFORMATION: "Xaa" at positions 38-40, 42-44, 46-47, 49-51, and 53-54 can be a
;   OTHER INFORMATION: ny amino acid
US-10-393-449-101

  Query Match             60.0%;  Score 63;  DB 15;  Length 91;
  Best Local Similarity   66.7%;  Pred. No. 0.85;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         65 AAKAAAEAAAKAAAEAAAKAAAAK 88


RESULT 45
US-10-177-725-47
; Sequence 47, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson, David
;  APPLICANT:  Bogenberger, Jakob M.
;  APPLICANT:  Peele, Beau R.
;  TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT SC
;  FILE REFERENCE: A-66900-4/RMS/AMS
;  CURRENT APPLICATION NUMBER: US/10/177,725
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: US 09/415,765
;  PRIOR FILING DATE: 1999-10-08
;  PRIOR APPLICATION NUMBER: US 09/169,015
;  PRIOR FILING DATE: 1998-10-08
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: synthetic
US-10-177-725-47

  Query Match             60.0%;  Score 63;  DB 14;  Length 104;
  Best Local Similarity   66.7%;  Pred. No. 0.98;
  Matches   16;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 AXAEAAEKAAKYAAEAAEKAAKAX 24
              |:  ||| ||| ||||| ||| |:
Db         79 AAKAAAEAAAKAAAEAAAKAAAAK 102


Search completed: July 11, 2005, 10:00:02
Job time : 160 secs
```

This Page Blank (uspto)